

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 13:45:48 ; Search time 2760 Seconds

(without alignments)
5883.828 Million cell updates/sec

Title: US-09-720-451-5

Perfect score: 1 tgattgtccaccattatt.....tcagtcatgcggccacgtg 558

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_da:*
2: gb_hlg:*
3: gb_hn:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------------|----------------------|
| 1 | 378.8 | 67.9 | 1897 | 8 | AF181967 | AF181967 Arabidops |
| 2 | 378.8 | 67.9 | 1917 | 8 | AY050434 | AY050434 Arabidops |
| 3 | 378.8 | 67.9 | 2034 | 8 | AE370515 | AE370515 Arabidops |
| 4 | 377.2 | 67.6 | 2037 | 8 | ATR245414 | ATR245414 Arabidops |
| 5 | 358.8 | 64.3 | 1810 | 8 | AY122922 | AY122922 Arabidops |
| 6 | 358.8 | 64.3 | 1994 | 8 | AF181966 | AF181966 Arabidops |
| 7 | 358.8 | 64.3 | 2002 | 8 | AY070034 | AY070034 Arabidops |
| 8 | 358.8 | 64.2 | 2225 | 8 | AF174486 | AF174486 Zea mays |
| 9 | 130 | 23.3 | 38571 | 3 | U39849 | U39849 Caenorhabdi |
| 10 | 127.2 | 22.8 | 1785 | 12 | AF441241 | AF441241 Synthetic |
| 11 | 126.8 | 22.7 | 81624 | 3 | AC099763 | AC099763 Caenorhab |
| 12 | 121.8 | 21.8 | 109741 | 8 | AC004005 | AC004005 S.pombe chr |
| 13 | 117.4 | 21.0 | 37093 | 8 | SPAC56F8 | SPAC56F8 Sequence |
| 14 | 116.4 | 20.9 | 1971 | 6 | AX069359 | AX069359 Sequence |
| 15 | 116.4 | 20.9 | 2187 | 6 | AR144956 | AR144956 Sequence |
| 16 | 116.4 | 20.9 | 2187 | 6 | AX050441 | AX050441 Sequence |
| 17 | 116.4 | 20.9 | 2196 | 12 | HSU09806 | U09806 Synthetic c |
| 18 | 116.4 | 20.9 | 2219 | 6 | A47328 | A47328 Sequence 3 |
| 19 | 116.4 | 20.9 | 2219 | 6 | AR097969 | AR097969 Sequence |
| 20 | 116.4 | 20.9 | 2219 | 6 | AR145811 | AR145811 Sequence |
| 21 | 116.4 | 20.9 | 2219 | 6 | AX356174 | AX356174 Sequence |
| 22 | 116.4 | 20.9 | 2219 | 6 | AX452351 | AX452351 Sequence |
| 23 | 116.4 | 20.9 | 2220 | 6 | A47326 | A47326 Sequence 1 |
| 24 | 116.4 | 20.9 | 2220 | 6 | AR097968 | AR097968 Sequence |
| 25 | 116.4 | 20.9 | 2220 | 6 | AR145810 | AR145810 Sequence |
| 26 | 116.4 | 20.9 | 2220 | 6 | AX356172 | AX356172 Sequence |
| 27 | 116.4 | 20.9 | 2313 | 9 | HS237672 | AJ237672 Homo sapi |
| 28 | 112.6 | 20.2 | 100906 | 8 | ATF24G16 | AL138647 Arabidops |
| 29 | 102.4 | 18.4 | 1250 | 10 | RNU57049 | U57049 Rattus norv |
| 30 | 67 | 12.0 | 3082 | 8 | SCYGL125W | Z72647 S.cerevisia |
| 31 | 67 | 12.0 | 3082 | 8 | SCSUP44GN | X94106 S.cerevisia |
| 32 | 64.6 | 11.6 | 110000 | 2 | LMFCCHR36_02 | Continuation (3 of |
| 33 | 56.2 | 10.1 | 15247 | 1 | AE000740 | AE000740 Aquifex a |
| 34 | 56.2 | 10.0 | 551 | 4 | AF239166 | AF239166 Sus scrof |
| 35 | 54 | 9.7 | 12141 | 1 | AE014080 | AE014080 Buchnera |
| 36 | 53.4 | 9.6 | 347550 | 1 | AP001118 | AP001118 Buchnera |
| 37 | 52.2 | 9.4 | 215050 | 1 | AL646057 | AL646057 Ralstonia |
| 38 | 51.8 | 9.3 | 198728 | 10 | MMETTER06 | AF105993 Mus muscu |
| 39 | 51.8 | 9.3 | 198728 | 10 | AL606929 | AL606929 Mouse DNA |
| 40 | 50.4 | 9.0 | 10592 | 1 | AE006058 | AE006058 Pasteurel |
| 41 | 50.2 | 9.0 | 181581 | 2 | AC113917 | AC113917 Rattus no |
| 42 | 49.2 | 8.8 | 10428 | 1 | AE006357 | AE006357 Lactococc |
| 43 | 48.8 | 8.7 | 11166 | 1 | U32823 | U32823 Haemophilus |
| 44 | 48.8 | 8.7 | 11166 | 1 | AX452358 | AX452358 Sequence |
| 45 | 48.6 | 8.7 | 235 | 6 | AX452358 | AX452358 Sequence |

ALIGNMENTS

RESULT 1
AF181967
LOCUS
DEFINITION
Arabidopsis thaliana methyltetrahydrofolate reductase MTHFR2
ACCESSION
AF181967
VERSION
AF181967.1 GI:5911426
KEYWORDS
SOURCE
Arabidopsis thaliana.
ORGANISM
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 1897)
Rojes, S., Wang, H., McNeill, S.D., Raymond, R.K., Appling, D.R.,

| TITLE | JOURNAL | MEDLINE | PUBMED | REFERENCE | AUTHORS | TITLE | JOURNAL | FEATURES | SOURCE |
|---|---|--------------------|-----------|---------------------|-----------------------------|-------------------|-------------------------|---|---------------------|
| Shachar-Hill, Y., Bohnett, H. J. and Hanson, A. D. | 7. Biol. Chem. 274 (51), 36089-36096 (1999) | 20062814 | 10593891 | 2 (bases 1 to 1897) | Roj'e, S. and Hanson, A. D. | Direct Submission | Submitted (31-AUG-1999) | Horticultural Sciences, University of Florida, Hall Road, Gainesville, FL 32611-0650, USA | Location/Qualifiers |
| Isolation, characterization, and functional expression of cDNAs encoding NADH-dependent methyltetrahydrofolate reductase from higher plants | | | | | | | | | |
| CDs | | | | | | | | | |
| 1. 1897 | | | | | | | | | |
| /EC.number="1.5.1.20" | | | | | | | | | |
| /function="catalyzes the reduction of 5,10-methylenetetrahydrofolate (CH2-THF) to 5-methyltetrahydrofolate (CH3-THF)" | | | | | | | | | |
| /note="NADH-dependent" | | | | | | | | | |
| /codon.start=1 | | | | | | | | | |
| /product="methyltetrahydrofolate reductase mtrfr2" | | | | | | | | | |
| /protein_id="AAD5788.1" | | | | | | | | | |
| /db_xref="GI:5911427" | | | | | | | | | |
| /translation="MKYIDKIOSLADGKTAPEFEPPKTEGVDNLFERRMDVAVY | | | | | | | | | |
| GETREDITGGGSGTADLLDIASRMNVVCSMMHLCTMPREKIDHLETRNS | | | | | | | | | |
| GIQNLAIKRGDPPGQKFEVQEGEDCALVDNLRSTIGYFEITVAGFEAPHV | | | | | | | | | |
| IGENLWASNEAYQSDLEKRIKIDAGADILVQLFPTDITGPNDCRQIGICPIV | | | | | | | | | |
| PEIMPIINNYRGFLKMTGFKTPIVEVMALIKIDNEEAVAYGILGTEMRKMLA | | | | | | | | | |
| HOVKSILHLYTLNMEKSAIILIMNIGMIDESKISRSIPMRPANYRTRKEDVRPFMAN | | | | | | | | | |
| RKRSYSIRTKGMEDEPQGRWSDRSASGYCALSDHSPRPAADKILQDEWVPLKSV | | | | | | | | | |
| DIQEKFEKLCGNLKSPPSELDGLQEPETIIMEDLIVNSGFLINSQSVAAERSG | | | | | | | | | |
| DSPTQKGGPGVGYOKAVLEFPCSKKIDAVYKEKRALSTITYAAVKKGRQVSNRA | | | | | | | | | |
| QADVAVWGVGFPAPKKEITQPIYDPAASFNVWDEAFETMSRMALYEPADPSRNLE | | | | | | | | | |
| EKNSYIVLSVENDYINGDIFRFPADL" | | | | | | | | | |
| BASE COUNT | 524 a | 362 c | 457 g | 354 t | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match | 67.9%; | Score 378.8; | DB 8; | Length 1897; | | | | | |
| Best Local Similarity | 79.9%; | Pred. No. 1.3e-85; | | | | | | | |
| Matches 446; | Conservative 0; | Mismatches 112; | Indels 0; | Gaps 0; | | | | | |
| QY 1 | TGATGTGACACCAATATTTATGTAGTACGAGATATATCCCAATTTGTGACAGACTGC | 60 | | | | | | | |
| DB 566 | TGATGTGACAGCTTTCCTATGATGATGATGATGATGATGATGATGATGATGATGATG | 625 | | | | | | | |
| QY 61 | GCCAAATTTGGAATACGTCCTATGTACCTGGAATATATGCCATTAATTAATTAACAG | 120 | | | | | | | |
| DB 626 | GGCMAATTTGGATTAATGTTGTCCTCATTTGTCCTGGAATATATGCCATTAATTAACAG | 685 | | | | | | | |
| QY 121 | GCTTATCCCGATGATGGGTTTGCCAAACAAGATATACAGGACATATATGGCGT | 180 | | | | | | | |
| DB 686 | GCTTTTGGCTGATGATGCTTTTCTGCAAGACTAAGATATACAGTTGAGGTATGGCTGC | 745 | | | | | | | |
| QY 181 | TAGAGCCTATCAAGACAAATGAAGAGCTGTCAAGGCTTATGAAATTCACCTGGGAATG | 240 | | | | | | | |
| DB 746 | TGGAGCCTATCAAGAGTAACGAAGAGCTGTGAAGAGCCTATGATTAACCTGGGAAG | 805 | | | | | | | |
| QY 241 | AAATGTCACAAAGATTTAGTCACTGATTAAGACTGCACTTATATACCTTAATAA | 300 | | | | | | | |
| DB 806 | AGATGTGTAAAAAGATGTTGGCTCTATGAGTCAAGCTCTTCATCTTATACATTAAACA | 865 | | | | | | | |
| QY 301 | TGGAGAATTCGCAATTTGGCAATACTAATGAACCTTGAGCTTAATTAAGAGCTCCAAATTT | 360 | | | | | | | |
| DB 866 | TGGAGAATTCGCTTTCGCAATTTGATGAATCTTGGATGATGATGATGATGATGATGATG | 925 | | | | | | | |
| QY 361 | CTAGGTCCTTATGAGAGAGCCCTGCAATGTTTCCGGTAAAGAGATGTCGCGC | 420 | | | | | | | |
| DB 926 | CTGCTTCTTATACCTTGAGAGAGCCCTGCAATGTTTTCCTGATTAAGAGAGATGTCGCGC | 985 | | | | | | | |
| QY 421 | CAATCTTTTGGCAATTCAGCAAAAGCCTCATATCAAGAGACCATGATGGATCAAT | 480 | | | | | | | |

| | TITLE | JOURNAL REFERENCE | AUTHORS | COMMENT |
|-------------------|--|---|---------|-----------------|
| D6 | 986 | CAATTTCTGTCAGCAACCGCCAAAGAGATCATTTCTTAGAACCAAGCGGTGGAGAATTC | 1045 | |
| OY | 481 | ACCACATGGGGCTTGGGGTGATTCCTGTATCCATCATATATGTCATTATTCGATTATC | 540 | |
| D6 | 1046 | TCCCACAAGGCCGGTGGGGGTGATTCACGAGATGCTTCATATGATGGACTCTGCATATCAC | 1105 | |
| OY | 541 | AGTTCACGCGGCCACGCTG | 558 | |
| D6 | 1106 | AGTTCCTACGCTCGCGAG | 1123 | |
| RESULT 2 | | | | |
| AY050434 | | 1917 bp | mRNA | linear |
| LOCUS | | | | pin 20-AUG-2001 |
| DEFINITION | Arabidopsis thaliana At2g44160/FelJ.29 mRNA, complete cds. | | | |
| ACCESSION | AY050434 | | | |
| VERSION | AY050434.1 | GI:15215809 | | |
| KEYWORDS | FLI CDNA. | | | |
| SOURCE | Arabidopsis thaliana. | | | |
| ORGANISM | Arabidopsis thaliana | | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. | | | |
| AUTHORS | 1 (bases 1 to 1917) | | | |
| TITLE | Shimn,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashitaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu.G., Yu.S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R. | | | |
| JOURNAL REFERENCE | Arabidopsis cDNA clones unpublished | | | |
| AUTHORS | 2 (bases 1 to 1917) | | | |
| TITLE | Shimn,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashitaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu.G., Yu.S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R. | | | |
| JOURNAL REFERENCE | Direct Submission Submitted (02-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA | | | |
| COMMENT | Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CNAs (RAFL CNA : 'RIKEN Arabidopsis Full-length cDNA' : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashitaki,Y. and Shinozaki,K.). | | | |
| FEATURES | The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shimn,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu.G., Yu.S., Davis,R.W., Theologis,A., and Ecker,J.R. | | | |
| SOURCE | Shimn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs. Location/Qualifiers 1..1917 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="2" /clone="RAFL05-16-121(RL3348)" | | | |

```
/note="ecotype: Columbia"
1..40
41..1825
/note="putative methylenetetrahydrofolate reductase"
/codon_start=1
/product="At2g44160/F6E13.29"
/protein_id="AAK91450.1"
/db_xref="GI:15215810"
/translation="MKVIDKIOSLADDEKTAFFSEFPFKTEDVDNLFERMDRAVAY
GPFCDITWAGSTADLTLDIASRMNVVCSMMHLCTMPEVEKIDHALETIRSN
GIQWVLALRGDPHGDQKRVYEGGFDALDVLNIRSKYGDYITVAGYPEABPDY
IGENGLASNEAYOSDLEYLKKRIDAGADLIVQLFYDIDFLKFNDCROJISCPYI
PCIMPLNNRGFLRMTEGCKTKIPEVMAALEPIKDNEEAVKAGIHLGTECKKMLA
HGKSLHLYTLNMEKSAIILNIGMIDESKISLPPWRPANNVFRKEEDVRIEMAN
RKYSITSRTKWEDEDPQGRWGSRSASGALSDFOSRPARADKLOQEWVPLKSE
DIOEKFEKIDGNLKSSPWSELDGLQEPETRIINEDLIVNSKGLTINSQPSVNAERS
DSPTVAGWGPVGYOKAYLEFFCSKEKIDAVEKCKALPSITYMAVNKGQWVSNIA
QADVNAVWGPVPAKEIIOPTIVDPASFVWMDAEFTWSRSMANLYPEADPSRNILE
EVKNSYILVSLVENDYINGIDIPAVFADL"
1826..1917
3'UTR
BASE COUNT      525 a      368 c      460 g      564 t
ORIGIN
Query Match      67.9%; Score 378.8; DB 8; Length 1917;
Best Local Similarity 79.9%; Pred. No. 1.3e-85;
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 1 TGATTGTACCCCAATTATTTATGATACGGAATATTCCTCAAAATTTGGAACGACTGTC 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 606 TGATTGTACTACGCTTTATGATACGATATATTCCTCAAGTTTGGAATGATGATG 655
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GCCAAATTTGGAATACGTCCTTATGTCCTGGAATTAAGCCATTAAATTAATGAAG 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 666 GCGAAATTTGGATTAATGTTGTCCTCATTTGCTTGGAATTAATTAATTAATGAG 725
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GCTTTATCCGCAATGACTGGGTTTGCAAAACAAGATACGAGTACATTATGCTGCTT 180
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 726 GCTTTTGGCTATGACTGGTTCTTGCAGACATTAACAGTGTAGGATGGCTGCCCT 785
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 TAGAGCTTATCAAGACATGAAGAAGCTGTCAAGCTTATGGAATTCACCTGGGACTG 240
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 786 TGAGGCTTATCAAGATTAACGAAGAAGCTGTGAAGCCCTATGCTTACCTTGAACAG 845
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 AAATGTCCAAAAGATTTTATGCTCATGGAATTAAGACATTCCTTTATACCTAATAA 300
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 846 AGATGTGTAAAAAGATGTTGGCTCATGAGTCAAGTCTTCTCATCTCTACACTTGAACA 905
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 TGGAGAAATCTGCATTGGCAATGCAATCACTTGCCCTAATTGAAGATCCAAAGTTT 360
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 906 TGGAGAAATCTGCTTTCGAATATGATGATGCTGTATGATGATGATGATGATGATG 965
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 CTAGGTCCTTACTTGGAGAGCGCCCTGCAAAATGTTTCCGCTGTTAAAGAAGATGTCGTC 420
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 966 CTGCTCTTATCCCTGTGAGAGCGCCCTGCAAAATGTTTCCGCTACTAAGGAAGATGTGCGCC 1025
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 CAATCTTTTGGGCAAAATCGAACAAAAGCTACATATCAAGAGCAATAGAGTGGGATCAAT 480
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1026 CCAATTTTCTGGCAAAACGCGCAAAAGCTACATTTCTTAAGACCAAGGCGTGGGAAGATT 1085
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 ACCACATGGGGCTTGGGGTGATTCCTGTATTCATCATATATGGTGCATTATCTGATTATTC 540
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1086 TCCACACAGGCGCGGTGGGATTCACGCGTCTTCATATATGATGATGATGATGATGATG 1145
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 AGTTCAATGGGCGCAGG 558
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1146 AGTTTCACGTCGCGAG 1163
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 3
AF370515 2034 bp mRNA linear PIN 30-APR-2001
LOCUS
DEFINITION Arabidopsis thaliana putative methylenetetrahydrofolate reductase
(AT2g44160; F6E13.29) mRNA, complete cds.
```

```
ACCESSION AF370515
VERSION AF370515.1 GI:13877628
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 2034)
REFERENCE
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sekurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT
e-mail for correspondence: arab@sequence.stanford.edu
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sekurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Southwick,A.,
Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.
Southwick,A. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.
FEATURES
source
location/Qualifiers
1..2034
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="11"
/clone="RAFL11-04-D03"
/note="this clone is in pb1uescript
ecotype: Columbia"
1..2034
/gene="At2g44160; F6E13.29"
91..1875
/gene="At2g44160; F6E13.29"
/codon_start=1
/product="putative methylenetetrahydrofolate reductase"
/protein_id="AAK3892.1"
/db_xref="GI:13877629"
/translation="MKVIDKIOSLADDEKTAFFSEFPFKTEDVDNLFERMDRAVAY
GPFCDITWAGSTADLTLDIASRMNVVCSMMHLCTMPEVEKIDHALETIRSN
GIQWVLALRGDPHGDQKRVYEGGFDALDVLNIRSKYGDYITVAGYPEABPDY
IGENGLASNEAYOSDLEYLKKRIDAGADLIVQLFYDIDFLKFNDCROJISCPYI
PCIMPLNNRGFLRMTEGCKTKIPEVMAALEPIKDNEEAVKAGIHLGTECKKMLA
HGKSLHLYTLNMEKSAIILNIGMIDESKISLPPWRPANNVFRKEEDVRIEMAN
RKYSITSRTKWEDEDPQGRWGSRSASGALSDFOSRPARADKLOQEWVPLKSE
DIOEKFEKIDGNLKSSPWSELDGLQEPETRIINEDLIVNSKGLTINSQPSVNAERS
DSPTVAGWGPVGYOKAYLEFFCSKEKIDAVEKCKALPSITYMAVNKGQWVSNIA
QADVNAVWGPVPAKEIIOPTIVDPASFVWMDAEFTWSRSMANLYPEADPSRNILE
EVKNSYILVSLVENDYINGIDIPAVFADL"
BASE COUNT 568 a 386 c 483 g 597 t
ORIGIN
Query Match 67.9%; Score 378.8; DB 8; Length 2034;
Best Local Similarity 79.9%; Pred. No. 1.3e-85;
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 1 TGATTGTACCCCAATTATTTATGATACGGAATATTCCTCAAAATTTGGAACGACTGTC 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 656 TGATTGTACTGAGCTTTTCTATGATGATGATATATTCCTCAAGTTTGTGAATGATGTC 715
QY 61 GCCAAATGGAAATACGTCGCTATTTGACCTGGAATTTGGCCATTTAATTAATCAAG 120
Db 716 GCGAAATGGGATTTACTTGTCCATTTCTCTGGAATTAAGCTTATTAATCAAGAG 775
QY 121 GCTTATTCGCAAGTACTGGGTTTTCGAAACAAAGATACAGCTGACATTAATGCTGCTT 180
Db 776 GCTTTTGGGATGACTGTTTCTGCAAGACTAAGATACCACTTGAAGTGAAGTGCCT 835
QY 181 TAGAGCTTCAAGGACATGAAGAAGCTGTGAAGCTTATGGAATTCACCTGGGAAGT 240
Db 836 TGAGGCTTCAAGGATTAAGGAAGAGCTGTGAAGGCTTATGATTAATCACTTGAAGAG 895
QY 241 AATGTGCAAAAAGATTTAGCTGATGATTAAGATTCATGATTCATTAATCACTAATA 300
Db 896 AGATGTGTAAGAAAGATTTGGCTCATGAGATCAAGTCTTATCTTACACATTTGAACA 955
QY 301 TGGAGAAATCTGCAATGGAATTAAGTATGACCTTGGCTTAATGAAGATCCAAAGTTT 360
Db 956 TGGAGAAATCTGCTCTTGCATATGATGATGATGATGATGATGATGATGATGATGAT 1015
QY 361 CTAAGGCTTACCTGGAAGAGGCTGCAATGTTTCCGTGTAAGAAGATGCTGCTC 420
Db 1016 CTGCTTCTTACCTGGAAGAGGCTGCAATGTTTCCGTGTAAGAAGATGCTGCTC 1075
QY 421 CAATCTTTTGGCAATCGAACCAAAAGCTACATATCAAGGACCAATGAGATGGATCAAT 480
Db 1076 CCATTTTCTGGGCAAAAGCTGCAAAAGATTTCTTCAAGGACCAAGGCTGGAAGATT 1135
QY 481 ACCCATGAGGCGTTGGGCTGATTCCTGTAATCAATGATGATGATGATGATGATGATG 540
Db 1136 TCCCAAGAGGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1195
QY 541 AGTTATCGGCGCAGCTG 558
Db 1196 AGTTCTCAGCTCCGCGAG 1213

RESULT 4
ATH245414
LOCUS 2037 bp mRNA linear PLN 01-SEP-1999
DEFINITION Arabidopsis thaliana mRNA for methylenetetrahydrofolate reductase.
ACCESSION AJ245414.1 GI:5823582
VERSION 5.10-methylenetetrahydrofolate reductase; methylenetetrahydrofolate
KEYWORDS reductase; MTHFR1 gene.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2037)
Ravanel, S., Rebelle, F. and Douce, R.
Folate metabolism in higher plants: cloning of a cDNA for
5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana
Unpublished
AUTHORS Ravel, S.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1999) Ravanel S., DBMS / PCV, URA-CNRS 576,
CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, FRANCE
FEATURES
source
1. 2037
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
68.1852
/gene="MTHFR1"
68.1852
/gene="MTHFR1"
/EC_number="1.5.1.20"
/function="reduction of methylenetetrahydrofolate to
methylenetetrahydrofolate"

/codon_start=1
/product="methylenetetrahydrofolate reductase"
/protein_id="CAB53783.1"
/db_xref="GI:5823583"
/translation="MKVVDKIQSLADEGKTAFSEFFPPKTEGDVNLPERMDRWAV
GPFCDITWAGSGSTADLTLDIASRMQSVLVESMMHLCSTNMPVEKIDHALETBSN
GIONVALRDPDHPGODKFOVQGFCDALDLYNHRSKYGDYFGITVAGYEAPDDY
IGBNGLASNEAYOSDLEYLKKRIDAGADI.LYQLFPDPIPLKFPVNDQDQISCPY
PGIMPINNRYGFLRMTEFCKTKLPVYMAALEPIKIDNEAVATKGIHIGTEKRMILA
HGKSLHLTYLNLNEKSALALIMNLGIDSKSLSRLPMRRPANVETKEDVPIFWAN
RPSYSISRTKGWDEPQGRGDSRSAYALSDHSPRBDKLIODEWVPLKVE
DIOEKFEKLEIDNLKSPSELDGLDPEKRIINLEOLIKVSKGFLITINSPVSNRS
DSPTVGMGPGVGVYOKAYLEFCSKEKIDAVYEKALPSITVYAVNGEOWSVTV
QADVNAVTVGVPFAKEIIOPTVIDPASFPVWMDENAEETSRSMANLYPEADSRNMLE
EVANSYLYSVLNDYINDGIDPVPDLD"

BASE COUNT 583 a 385 c 480 g 589 t
ORIGIN

Query Match 67.6%; Score 377.2; DB 8; Length 2037;
Best Local Similarity 79.7%; Pred. No. 3,2e-85;
Matches 445; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 1 TGATTGTACCCCAATTAATTTTANGATAGGATATATTCCTCAAAATTTGGAGAGCTGTC 60
Db 633 TGATTGTACTACACCTTTCTATGATGATGATGATGATGATGATGATGATGATGATG 692
QY 61 GCCAAATGGAAATACGTCGCTATTTGACCTGGAATTAAGCCATTAATTAATCAAG 120
Db 693 GCGAAATTTGGATTAAGTTGTCCTGCTGCTGGAATTAAGCTTATTAATCAATCAAG 752
QY 121 GCTTATTCGCAATGACTGGGTTTTCGAAACAAAGATACCGAGCTGATTAATGCTGCTT 180
Db 753 GCTTCTTGGCATGACTGTTTTCGCAAGACTAAGATACCAAGTTGAGGATGCTGCTC 812
QY 181 TAGAGCTTCAAGGACAAATGAAGAGCTGTCAAGCTTATGGAATTTACCTGGGAAGT 240
Db 813 TGGAGCTTATCAAGGATTAAGGAAGAGCTGTGAAGGCTTGTGATTTACCTTGGAGAG 872
QY 241 AATGTGCAAAAAGATTTAGCTGATGATTAAGATTCATGATTCATTTATGACTAATA 300
Db 873 AGATGTGTAAGAAAGATTTGCTCTCATGAGACTCAAGTCTTCTATCTTACACATTAACA 932
QY 301 TGGAGAAATCTGCAATGGAATTAAGTATGACCTTGGCTTAATGAAGATCCAAAGTTT 360
Db 933 TGGAGAAATCTGCTCTTGCATTAATGATGATGATGATGATGATGATGATGATGATG 992
QY 361 CTAGGCTTACCTTGGAGAGCGCCGTCGAATGTTTCCGTGTTAAGAAGATGCTCCGC 420
Db 993 CTGCTTCTTACCTTGGAGAGCGCCGTCGAATGTTTCCGTGTTAAGAAGATGCTCCGC 1052
QY 421 CAATCTTTTGGCAAAATCGAACCAAAAGCTACATTAAGGACCAATGAGATGGATCAAT 480
Db 1053 CCATTTTCTGGCAAAAGCTGTCGAAGAGCTATTTCTTAAGACCAAGGCTGGGAAGACT 1112
QY 481 ACCCAATGAGGCGTTGGGCTGATTCCTGTAATCCATCATATGCTGATTAATGATTAATC 540
Db 1113 TCCCAAGAGCGGTTGGGCTGATTAAGAGCTGCTTCAATATGATGATGATGATGATG 1172
QY 541 AGTTATCGGCGCAGCTG 558
Db 1173 AGTTCTCAGCTCCGCGAG 1190

RESULT 5
AY122922
LOCUS 1810 bp mRNA linear PLN 03-JUL-2002
DEFINITION Arabidopsis thaliana putative methylenetetrahydrofolate reductase
MTHFR1 (At3g59970) mRNA, complete cds.
ACCESSION AY122922
VERSION AY122922.1 GI:21689666
KEYWORDS FLI-cDNA.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1810)

TITLE
JOURNAL
REFERENCE
AUTHORS

Yamada, K., Chan, M.-M., Chang, C.-H., Dale, J.-M., Deng, J.-M., Hsuan, Y.-W., Lee, J.-M., Quach, H.-L., Tang, C.-C., Toriumi, M., Wu, H.-C., Yu, G., Bower, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.-J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Arabidopsis Open Reading Frame (ORF) Clones Unpublished

TITLE
JOURNAL
COMMENT

2 (bases 1 to 1810)
Yamada, K., Chan, M.-M., Chang, C.-H., Dale, J.-M., Deng, J.-M., Hsuan, Y.-W., Lee, J.-M., Quach, H.-L., Tang, C.-C., Toriumi, M., Wu, H.-C., Yu, G., Bower, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.-J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Direct Submission Submitted (17-JUN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

FEATURES

source

The Saik, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Chan, M.-M., Chang, C.-H., Dale, J.-M., Deng, J.-M., Hsuan, Y.-W., Lee, J.-M., Quach, H.-L., Tang, C.-C., Toriumi, M., Wu, H.-C., Yu, G., Bower, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.-J., Shin, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Location/Qualifiers

gene
CDS

1..1810
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U18407"
/note="This clone is in PUNI 51.
ecotype: Columbia"
1..1810
/gene="At3g59970"
1..1779
/gene="At3g59970"
/codon_start=1
/evidence=experimental
/product="putative methylentetrahydrofolate reductase MTHFR1"

/protein_id="AA067455.1"

/db_xref="GI:2168967"

/translation="MKVVDKIKSVTEGQATSEFPFPPKTEGVEMLFERMDRLVY
GPTFCIDITWAGAGSTADLTLETASRMQVAVCEVMTKCTNMPEKIDHAELETIRSN
GIONVIALRGDPHGGDFOVGEFGAALDVNIIRSKYDGYFAGVPEAHPDV
TEADGLATPEYQSDLAIVKKKYNAGADLTITOLPYDIDFLKYNOCROGICNPY
PGMPSINVKGLRMAGFCKTIRPARELTALAEPIKNDENAVKATGIRHATCMCKKILA
HGITSILHTLVNDSKAIGILMNLGLIDESKISRLEPWRRAVNFRIKEDVAPLEMAN
RPKYSIRHTKWNDEPHGRWDSHAAYSTLSDYOFAPRKGRDKLQDEVVPLKSI
DVOEKFEKICLNLSKSPWSEIDGQPEPTKILINQDGIKINSNGEFLIKSSQVNAKS
DSPAIQMGPGGYYOKAYLEFFGSKDKLDLTVKSKAFPSITTAIVKSNWVNSNG
ESDVNAVTVGWVPAKEVLOPTITVDPASFKWKDEAFETIWSMANLTYEDDPCKILLE
EYKNSYIYSLVDNNYINGDIFSVA"

1780..1810

/gene="At3g59970"

508 a 372 c 427 g 503 t

3' UTR

BASE COUNT

ORIGIN

Query Match 64.3%; Score 358.8; DB 8; Length 1810;
Best Local Similarity 78.0%; Pred. No. 1.5e-80;
Matches 432; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 1 TGATTGCAACCAATTTATTTATGATACGATATATCCCAATTTGCAAGACGTC 60
DB 566 TGATTGTACACACCTTTCTATGATCTGATATATTTCTCAAGTTGTAAAGATGTC 625
QY 61 GCCAAATGGAATTAACCTGTCTTATTTACCTGGAATTAATTAATTAACAG 120
DB 626 GGCAAATCGGATTAATTTGTCCTCATTTGTCCTGGAATTAATTAATTAACAG 685
QY 121 GCTTATCCGATGACCTGCTTTGCAAAACAAGATTAACGCTGACATTAAGCTCTT 180
DB 686 GGTTCCTGCGATGCTGCTGCTTCTGTAAGACCAAGATTAACGCTGACCTGCTCT 745
QY 181 TAGAGCCATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
DB 746 TAGAGCCATTAAGGATTAATGACGAGGCTGTTAAGCCTTAAGATTAATTAATTA 805
QY 241 AATGTGCAAAAGATTTTACCTCATGGAATTAAGACATTTGATTAACATAATA 300
DB 806 AATGTGCAAAAGATTTTGGCCCATGGAATTAATTAATTAATTAATTAATTA 865
QY 301 TGGAGAAATCTGCAATGGAATTAATTAATTAATTAATTAATTAATTAATTA 360
DB 866 TGGAGAAATCTGCAATGGAATTAATTAATTAATTAATTAATTAATTAATTA 925
QY 361 CTAGGCTCTTACCTTGGAGACGCGCTGCAAAATGTTTCCGTTAAGAGATGTCGTC 420
DB 926 CTGCTCTCTACCTTGGAGACGCGCTGCAAAATGTTTCCGTTAAGAGATGTCGTC 985
QY 421 CAATCTTTGGGCAATGCAACCAAAAGATTAATTAATTAATTAATTAATTAATTA 480
DB 986 CAATCTTTGGGCAATGCAACCAAAAGATTAATTAATTAATTAATTAATTAATTA 1045
QY 481 ACCACATGAGGCGCTTGGGCTGATTCCTGTAATCATATATGATGCAATTCGATTA 540
DB 1046 TCCACATGAGGCGCTTGGGCTGATTCACACAGTACAGTACAGTACAGTACAG 1105
QY 541 AGTTCATGCGGCA 554
DB 1106 AGTTCGCGGCA 1119
RESULT 6
AF181966
LOCUS
DEFINITION
Arabidopsis thaliana methylentetrahydrofolate reductase MTHFR1
ACCESSION
AF181966
VERSION
AF181966.1 GI:5911424
KEYWORDS
ORIGIN
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424
Arabidopsis thaliana.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots, Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 1994)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1994 bp mRNA linear PLN 13-DEC-1999
mRNA, complete cds.
AF181966.1 GI:5911424

YK1335b02.3"
/codon_start=1
/product="Hypothetical protein C06A8.8"
/protein_id="AA81046.2"
/db_xref="GI:21913123"
/translation="MPLPFSIIISTTSQDLVDADSDDEAEFTYVNMVITERNHSPQSM
VLELEESLAVSSVKLIVKRDYPPSSLTAVSLGIPSHFNRSRGSYNADYVROELSPR
DDEHVEYQGRHKPAKVEFNSPAKPLIMNLFKISIVHGSDDVDNHOYSILSVIS
GAGEILPDSHLKDEPKQTEMYSIELDGIILPKSRKMEETIYGLVDTPDSEDE
ILFSKMLNRKREAVMKNPOLPTTSATLETIALDSCENSAVSSSSSKQSSVS
PFSDR"
9378. .10828
/gene="C06A8.6"
/note="for a graphical representation of this gene see:
http://www.wormbase.org/db/seq/sequence?name=C06A8.6;class
=sequence"
join(9578. .9664,9711. .10229,10279. .10461,10518. .10676,
10727. .10828)"
/gene="C06A8.6"
/note="similar to leucine-rich repeat regions of L.
monocytogenes Internalin and S. pombe Sds22"
/product="Hypothetical protein C06A8.6"
/codon_start=1
/protein_id="AA81047.1"
/db_xref="GI:1055044"
/translation="MNYTSVSESGVKFEKIVPEPDIETVEIPAKAVDILRHRI
KEIGDYSLWTHVHESFRNWLKIKTENIDCITPLHFEITNOITKVENDSLVNES
LDLSEFNKIKTEMLERLTKLTFVHKTKIKISBLDTLELELGDNRKIAEML
DNMLKIDRLFLGANOIRLLENVDHLKLVLSLPANAITYVDNISGLHMKIEYLAQ
GIRYVCGIDENHLELIDFNQNRLEKVENIHQKILDFNRAEGRKLDNWSIMDELVL
POLNSYVDENPIASVDYRGRKVIHLDOITRLDQTCREHSMLRPRKEKNETELSR
VLDNPSDHPRESY"
complement(11104. .14712)
/gene="C06A8.1"
/note="for a graphical representation of this gene see:
http://www.wormbase.org/db/seq/sequence?name=C06A8.1a;class
=sequence"
complement(join(11104. .11257,11302. .11456,11506. .11919,
11967. .12704,12762. .13134,13185. .13259,14630. .14712))
/gene="C06A8.1"
/note="similar to 5,10-methyltetrahydrofolate
reductases; coded for by the following C. elegans cDNAs:
cm13h3, YK4d8.3, YK4d8.5, YK5a2.3, YK5a2.5, YK10a4.3,
YK10a4.5, YK11e1.5, YK11e2.5, YK11e3.3, YK11e4.5,
YK11e5.3, YK11e6.3, YK11e7.3, YK11e8.3, YK11e9.3,
YK11e10.3, YK11e11.3, YK11e12.3, YK11e13.3, YK11e14.3,
YK11e15.3, YK11e16.3, YK11e17.3, YK11e18.3, YK11e19.3,
YK11e20.3, YK11e21.3, YK11e22.3, YK11e23.3, YK11e24.3,
YK11e25.3, YK11e26.3, YK11e27.3, YK11e28.3, YK11e29.3,
YK11e30.3, YK11e31.3, YK11e32.3, YK11e33.3, YK11e34.3,
YK11e35.3, YK11e36.3, YK11e37.3, YK11e38.3, YK11e39.3,
YK11e40.3, YK11e41.3, YK11e42.3, YK11e43.3, YK11e44.3,
YK11e45.3, YK11e46.3, YK11e47.3, YK11e48.3, YK11e49.3,
YK11e50.3, YK11e51.3, YK11e52.3, YK11e53.3, YK11e54.3,
YK11e55.3, YK11e56.3, YK11e57.3, YK11e58.3, YK11e59.3,
YK11e60.3, YK11e61.3, YK11e62.3, YK11e63.3, YK11e64.3,
YK11e65.3, YK11e66.3, YK11e67.3, YK11e68.3, YK11e69.3,
YK11e70.3, YK11e71.3, YK11e72.3, YK11e73.3, YK11e74.3,
YK11e75.3, YK11e76.3, YK11e77.3, YK11e78.3, YK11e79.3,
YK11e80.3, YK11e81.3, YK11e82.3, YK11e83.3, YK11e84.3,
YK11e85.3, YK11e86.3, YK11e87.3, YK11e88.3, YK11e89.3,
YK11e90.3, YK11e91.3, YK11e92.3, YK11e93.3, YK11e94.3,
YK11e95.3, YK11e96.3, YK11e97.3, YK11e98.3, YK11e99.3,
YK11e100.3, YK11e101.3, YK11e102.3, YK11e103.3, YK11e104.3,
YK11e105.3, YK11e106.3, YK11e107.3, YK11e108.3, YK11e109.3,
YK11e110.3, YK11e111.3, YK11e112.3, YK11e113.3, YK11e114.3,
YK11e115.3, YK11e116.3, YK11e117.3, YK11e118.3, YK11e119.3,
YK11e120.3, YK11e121.3, YK11e122.3, YK11e123.3, YK11e124.3,
YK11e125.3, YK11e126.3, YK11e127.3, YK11e128.3, YK11e129.3,
YK11e130.3, YK11e131.3, YK11e132.3, YK11e133.3, YK11e134.3,
YK11e135.3, YK11e136.3, YK11e137.3, YK11e138.3, YK11e139.3,
YK11e140.3, YK11e141.3, YK11e142.3, YK11e143.3, YK11e144.3,
YK11e145.3, YK11e146.3, YK11e147.3, YK11e148.3, YK11e149.3,
YK11e150.3, YK11e151.3, YK11e152.3, YK11e153.3, YK11e154.3,
YK11e155.3, YK11e156.3, YK11e157.3, YK11e158.3, YK11e159.3,
YK11e160.3, YK11e161.3, YK11e162.3, YK11e163.3, YK11e164.3,
YK11e165.3, YK11e166.3, YK11e167.3, YK11e168.3, YK11e169.3,
YK11e170.3, YK11e171.3, YK11e172.3, YK11e173.3, YK11e174.3,
YK11e175.3, YK11e176.3, YK11e177.3, YK11e178.3, YK11e179.3,
YK11e180.3, YK11e181.3, YK11e182.3, YK11e183.3, YK11e184.3,
YK11e185.3, YK11e186.3, YK11e187.3, YK11e188.3, YK11e189.3,
YK11e190.3, YK11e191.3, YK11e192.3, YK11e193.3, YK11e194.3,
YK11e195.3, YK11e196.3, YK11e197.3, YK11e198.3, YK11e199.3,
YK11e200.3, YK11e201.3, YK11e202.3, YK11e203.3, YK11e204.3,
YK11e205.3, YK11e206.3, YK11e207.3, YK11e208.3, YK11e209.3,
YK11e210.3, YK11e211.3, YK11e212.3, YK11e213.3, YK11e214.3,
YK11e215.3, YK11e216.3, YK11e217.3, YK11e218.3, YK11e219.3,
YK11e220.3, YK11e221.3, YK11e222.3, YK11e223.3, YK11e224.3,
YK11e225.3, YK11e226.3, YK11e227.3, YK11e228.3, YK11e229.3,
YK11e230.3, YK11e231.3, YK11e232.3, YK11e233.3, YK11e234.3,
YK11e235.3, YK11e236.3, YK11e237.3, YK11e238.3, YK11e239.3,
YK11e240.3, YK11e241.3, YK11e242.3, YK11e243.3, YK11e244.3,
YK11e245.3, YK11e246.3, YK11e247.3, YK11e248.3, YK11e249.3,
YK11e250.3, YK11e251.3, YK11e252.3, YK11e253.3, YK11e254.3,
YK11e255.3, YK11e256.3, YK11e257.3, YK11e258.3, YK11e259.3,
YK11e260.3, YK11e261.3, YK11e262.3, YK11e263.3, YK11e264.3,
YK11e265.3, YK11e266.3, YK11e267.3, YK11e268.3, YK11e269.3,
YK11e270.3, YK11e271.3, YK11e272.3, YK11e273.3, YK11e274.3,
YK11e275.3, YK11e276.3, YK11e277.3, YK11e278.3, YK11e279.3,
YK11e280.3, YK11e281.3, YK11e282.3, YK11e283.3, YK11e284.3,
YK11e285.3, YK11e286.3, YK11e287.3, YK11e288.3, YK11e289.3,
YK11e290.3, YK11e291.3, YK11e292.3, YK11e293.3, YK11e294.3,
YK11e295.3, YK11e296.3, YK11e297.3, YK11e298.3, YK11e299.3,
YK11e300.3, YK11e301.3, YK11e302.3, YK11e303.3, YK11e304.3,
YK11e305.3, YK11e306.3, YK11e307.3, YK11e308.3, YK11e309.3,
YK11e310.3, YK11e311.3, YK11e312.3, YK11e313.3, YK11e314.3,
YK11e315.3, YK11e316.3, YK11e317.3, YK11e318.3, YK11e319.3,
YK11e320.3, YK11e321.3, YK11e322.3, YK11e323.3, YK11e324.3,
YK11e325.3, YK11e326.3, YK11e327.3, YK11e328.3, YK11e329.3,
YK11e330.3, YK11e331.3, YK11e332.3, YK11e333.3, YK11e334.3,
YK11e335.3, YK11e336.3, YK11e337.3, YK11e338.3, YK11e339.3,
YK11e340.3, YK11e341.3, YK11e342.3, YK11e343.3, YK11e344.3,
YK11e345.3, YK11e346.3, YK11e347.3, YK11e348.3, YK11e349.3,
YK11e350.3, YK11e351.3, YK11e352.3, YK11e353.3, YK11e354.3,
YK11e355.3, YK11e356.3, YK11e357.3, YK11e358.3, YK11e359.3,
YK11e360.3, YK11e361.3, YK11e362.3, YK11e363.3, YK11e364.3,
YK11e365.3, YK11e366.3, YK11e367.3, YK11e368.3, YK11e369.3,
YK11e370.3, YK11e371.3, YK11e372.3, YK11e373.3, YK11e374.3,
YK11e375.3, YK11e376.3, YK11e377.3, YK11e378.3, YK11e379.3,
YK11e380.3, YK11e381.3, YK11e382.3, YK11e383.3, YK11e384.3,
YK11e385.3, YK11e386.3, YK11e387.3, YK11e388.3, YK11e389.3,
YK11e390.3, YK11e391.3, YK11e392.3, YK11e393.3, YK11e394.3,
YK11e395.3, YK11e396.3, YK11e397.3, YK11e398.3, YK11e399.3,
YK11e400.3, YK11e401.3, YK11e402.3, YK11e403.3, YK11e404.3,
YK11e405.3, YK11e406.3, YK11e407.3, YK11e408.3, YK11e409.3,
YK11e410.3, YK11e411.3, YK11e412.3, YK11e413.3, YK11e414.3,
YK11e415.3, YK11e416.3, YK11e417.3, YK11e418.3, YK11e419.3,
YK11e420.3, YK11e421.3, YK11e422.3, YK11e423.3, YK11e424.3,
YK11e425.3, YK11e426.3, YK11e427.3, YK11e428.3, YK11e429.3,
YK11e430.3, YK11e431.3, YK11e432.3, YK11e433.3, YK11e434.3,
YK11e435.3, YK11e436.3, YK11e437.3, YK11e438.3, YK11e439.3,
YK11e440.3, YK11e441.3, YK11e442.3, YK11e443.3, YK11e444.3,
YK11e445.3, YK11e446.3, YK11e447.3, YK11e448.3, YK11e449.3,
YK11e450.3, YK11e451.3, YK11e452.3, YK11e453.3, YK11e454.3,
YK11e455.3, YK11e456.3, YK11e457.3, YK11e458.3, YK11e459.3,
YK11e460.3, YK11e461.3, YK11e462.3, YK11e463.3, YK11e464.3,
YK11e465.3, YK11e466.3, YK11e467.3, YK11e468.3, YK11e469.3,
YK11e470.3, YK11e471.3, YK11e472.3, YK11e473.3, YK11e474.3,
YK11e475.3, YK11e476.3, YK11e477.3, YK11e478.3, YK11e479.3,
YK11e480.3, YK11e481.3, YK11e482.3, YK11e483.3, YK11e484.3,
YK11e485.3, YK11e486.3, YK11e487.3, YK11e488.3, YK11e489.3,
YK11e490.3, YK11e491.3, YK11e492.3, YK11e493.3, YK11e494.3,
YK11e495.3, YK11e496.3, YK11e497.3, YK11e498.3, YK11e499.3,
YK11e500.3, YK11e501.3, YK11e502.3, YK11e503.3, YK11e504.3,
YK11e505.3, YK11e506.3, YK11e507.3, YK11e508.3, YK11e509.3,
YK11e510.3, YK11e511.3, YK11e512.3, YK11e513.3, YK11e514.3,
YK11e515.3, YK11e516.3, YK11e517.3, YK11e518.3, YK11e519.3,
YK11e520.3, YK11e521.3, YK11e522.3, YK11e523.3, YK11e524.3,
YK11e525.3, YK11e526.3, YK11e527.3, YK11e528.3, YK11e529.3,
YK11e530.3, YK11e531.3, YK11e532.3, YK11e533.3, YK11e534.3,
YK11e535.3, YK11e536.3, YK11e537.3, YK11e538.3, YK11e539.3,
YK11e540.3, YK11e541.3, YK11e542.3, YK11e543.3, YK11e544.3,
YK11e545.3, YK11e546.3, YK11e547.3, YK11e548.3, YK11e549.3,
YK11e550.3, YK11e551.3, YK11e552.3, YK11e553.3, YK11e554.3,
YK11e555.3, YK11e556.3, YK11e557.3, YK11e558.3, YK11e559.3,
YK11e560.3, YK11e561.3, YK11e562.3, YK11e563.3, YK11e564.3,
YK11e565.3, YK11e566.3, YK11e567.3, YK11e568.3, YK11e569.3,
YK11e570.3, YK11e571.3, YK11e572.3, YK11e573.3, YK11e574.3,
YK11e575.3, YK11e576.3, YK11e577.3, YK11e578.3, YK11e579.3,
YK11e580.3, YK11e581.3, YK11e582.3, YK11e583.3, YK11e584.3,
YK11e585.3, YK11e586.3, YK11e587.3, YK11e588.3, YK11e589.3,
YK11e590.3, YK11e591.3, YK11e592.3, YK11e593.3, YK11e594.3,
YK11e595.3, YK11e596.3, YK11e597.3, YK11e598.3, YK11e599.3,
YK11e600.3, YK11e601.3, YK11e602.3, YK11e603.3, YK11e604.3,
YK11e605.3, YK11e606.3, YK11e607.3, YK11e608.3, YK11e609.3,
YK11e610.3, YK11e611.3, YK11e612.3, YK11e613.3, YK11e614.3,
YK11e615.3, YK11e616.3, YK11e617.3, YK11e618.3, YK11e619.3,
YK11e620.3, YK11e621.3, YK11e622.3, YK11e623.3, YK11e624.3,
YK11e625.3, YK11e626.3, YK11e627.3, YK11e628.3, YK11e629.3,
YK11e630.3, YK11e631.3, YK11e632.3, YK11e633.3, YK11e634.3,
YK11e635.3, YK11e636.3, YK11e637.3, YK11e638.3, YK11e639.3,
YK11e640.3, YK11e641.3, YK11e642.3, YK11e643.3, YK11e644.3,
YK11e645.3, YK11e646.3, YK11e647.3, YK11e648.3, YK11e649.3,
YK11e650.3, YK11e651.3, YK11e652.3, YK11e653.3, YK11e654.3,
YK11e655.3, YK11e656.3, YK11e657.3, YK11e658.3, YK11e659.3,
YK11e660.3, YK11e661.3, YK11e662.3, YK11e663.3, YK11e664.3,
YK11e665.3, YK11e666.3, YK11e667.3, YK11e668.3, YK11e669.3,
YK11e670.3, YK11e671.3, YK11e672.3, YK11e673.3, YK11e674.3,
YK11e675.3, YK11e676.3, YK11e677.3, YK11e678.3, YK11e679.3,
YK11e680.3, YK11e681.3, YK11e682.3, YK11e683.3, YK11e684.3,
YK11e685.3, YK11e686.3, YK11e687.3, YK11e688.3, YK11e689.3,
YK11e690.3, YK11e691.3, YK11e692.3, YK11e693.3, YK11e694.3,
YK11e695.3, YK11e696.3, YK11e697.3, YK11e698.3, YK11e699.3,
YK11e700.3, YK11e701.3, YK11e702.3, YK11e703.3, YK11e704.3,
YK11e705.3, YK11e706.3, YK11e707.3, YK11e708.3, YK11e709.3,
YK11e710.3, YK11e711.3, YK11e712.3, YK11e713.3, YK11e714.3,
YK11e715.3, YK11e716.3, YK11e717.3, YK11e718.3, YK11e719.3,
YK11e720.3, YK11e721.3, YK11e722.3, YK11e723.3, YK11e724.3,
YK11e725.3, YK11e726.3, YK11e727.3, YK11e728.3, YK11e729.3,
YK11e730.3, YK11e731.3, YK11e732.3, YK11e733.3, YK11e734.3,
YK11e735.3, YK11e736.3, YK11e737.3, YK11e738.3, YK11e739.3,
YK11e740.3, YK11e741.3, YK11e742.3, YK11e743.3, YK11e744.3,
YK11e745.3, YK11e746.3, YK11e747.3, YK11e748.3, YK11e749.3,
YK11e750.3, YK11e751.3, YK11e752.3, YK11e753.3, YK11e754.3,
YK11e755.3, YK11e756.3, YK11e757.3, YK11e758.3, YK11e759.3,
YK11e760.3, YK11e761.3, YK11e762.3, YK11e763.3, YK11e764.3,
YK11e765.3, YK11e766.3, YK11e767.3, YK11e768.3, YK11e769.3,
YK11e770.3, YK11e771.3, YK11e772.3, YK11e773.3, YK11e774.3,
YK11e775.3, YK11e776.3, YK11e777.3, YK11e778.3, YK11e779.3,
YK11e780.3, YK11e781.3, YK11e782.3, YK11e783.3, YK11e784.3,
YK11e785.3, YK11e786.3, YK11e787.3, YK11e788.3, YK11e789.3,
YK11e790.3, YK11e791.3, YK11e792.3, YK11e793.3, YK11e794.3,
YK11e795.3, YK11e796.3, YK11e797.3, YK11e798.3, YK11e799.3,
YK11e800.3, YK11e801.3, YK11e802.3, YK11e803.3, YK11e804.3,
YK11e805.3, YK11e806.3, YK11e807.3, YK11e808.3, YK11e809.3,
YK11e810.3, YK11e811.3, YK11e812.3, YK11e813.3, YK11e814.3,
YK11e815.3, YK11e816.3, YK11e817.3, YK11e818.3, YK11e819.3,
YK11e820.3, YK11e821.3, YK11e822.3, YK11e823.3, YK11e824.3,
YK11e825.3, YK11e826.3, YK11e827.3, YK11e828.3, YK11e829.3,
YK11e830.3, YK11e831.3, YK11e832.3, YK11e833.3, YK11e834.3,
YK11e835.3, YK11e836.3, YK11e837.3, YK11e838.3, YK11e839.3,
YK11e840.3, YK11e841.3, YK11e842.3, YK11e843.3, YK11e844.3,
YK11e845.3, YK11e846.3, YK11e847.3, YK11e848.3, YK11e849.3,
YK11e850.3, YK11e851.3, YK11e852.3, YK11e853.3, YK11e854.3,
YK11e855.3, YK11e856.3, YK11e857.3, YK11e858.3, YK11e859.3,
YK11e860.3, YK11e861.3, YK11e862.3, YK11e863.3, YK11e864.3,
YK11e865.3, YK11e866.3, YK11e867.3, YK11e868.3, YK11e869.3,
YK11e870.3, YK11e871.3, YK11e872.3, YK11e873.3, YK11e874.3,
YK11e875.3, YK11e876.3, YK11e877.3, YK11e878.3, YK11e879.3,
YK11e880.3, YK11e881.3, YK11e882.3, YK11e883.3, YK11e884.3,
YK11e885.3, YK11e886.3, YK11e887.3, YK11e888.3, YK11e889.3,
YK11e890.3, YK11e891.3, YK11e892.3, YK11e893.3, YK11e894.3,
YK11e895.3, YK11e896.3, YK11e897.3, YK11e898.3, YK11e899.3,
YK11e900.3, YK11e901.3, YK11e902.3, YK11e903.3, YK11e904.3,
YK11e905.3, YK11e906.3, YK11e907.3, YK11e908.3, YK11e909.3,
YK11e910.3, YK11e911.3, YK11e912.3, YK11e913.3, YK11e914.3,
YK11e915.3, YK11e916.3, YK11e917.3, YK11e918.3, YK11e919.3,
YK11e920.3, YK11e921.3, YK11e922.3, YK11e923.3, YK11e924.3,
YK11e925.3, YK11e926.3, YK11e927.3, YK11e928.3, YK11e929.3,
YK11e930.3, YK11e931.3, YK11e932.3, YK11e933.3, YK11e934.3,
YK11e935.3, YK11e936.3, YK11e937.3, YK11e938.3, YK11e939.3,
YK11e940.3, YK11e941.3, YK11e942.3, YK11e943.3, YK11e944.3,
YK11e945.3, YK11e946.3, YK11e947.3, YK11e948.3, YK11e949.3,
YK11e950.3, YK11e951.3, YK11e952.3, YK11e953.3, YK11e954.3,
YK11e955.3, YK11e956.3, YK11e957.3, YK11e958.3, YK11e959.3,
YK11e960.3, YK11e961.3, YK11e962.3, YK11e963.3, YK11e964.3,
YK11e965.3, YK11e966.3, YK11e967.3, YK11e968.3, YK11e969.3,
YK11e970.3, YK11e971.3, YK11e972.3, YK11e973.3, YK11e974.3,
YK11e975.3, YK11e976.3, YK11e977.3, YK11e978.3, YK11e979.3,
YK11e980.3, YK11e981.3, YK11e982.3, YK11e983.3, YK11e984.3,
YK11e985.3, YK11e986.3, YK11e987.3, YK11e988.3, YK11e989.3,
YK11e990.3, YK11e991.3, YK11e992.3, YK11e993.3, YK11e994.3,
YK11e995.3, YK11e996.3, YK11e997.3, YK11e9

| | | | |
|----|------|---|------|
| Db | 7102 | AGGGAAGCATTCGTGAGATTTAAAGCTCTTGATTTGTGGAAGCTCGAAGGAGATCGTG | 7043 |
| Qy | 364 | GGTCCCTACCTTGGAGACGCCCTGCAAAATGTTTTCCGTGTTAAAGAAGATGTCCTGCCAA | 423 |
| Db | 7042 | TGTTCCCATGAAAAATCGTTCTCAACATCCCAATCAGATGTCTTGAAAGGTTGCCCAA | 6983 |
| Qy | 424 | TCTTTTGGGCAATCGACCAAAACCTATCATATCAAGGACCATAGATGGATCAATAC | 483 |
| Db | 6982 | TCTACGTGAGACTTCGCTCCGCACTTATATACACAGAACTGGGATTTGGATCAATTC | 6923 |
| Qy | 484 | CACATGGGCGTTTGGGCGATTCCCTAATCATCATATGGTG | 525 |
| Db | 6922 | CGAAGCGAATATGGGAAACAGTTCTTCTCCAGATTTGGTG | 6881 |

| RESULT 12 | AC004005 | LOCUS | DEFINITION |
|-----------|----------|-----------|---|
| | AC004005 | 109741 bp | DNA |
| | | | linear |
| | | | PLN 11-MAR-2002 |
| | | | Arabidopsis thaliana chromosome 2 clone FGE13 map |
| | | | C1C10F02, C1C02E07, complete sequence. |

| repeat_region | complement(5593. .5613) |
|---------------|---|
| repeat_region | /rpt_family="At_rich" |
| repeat_region | complement(6730. 6753) |
| repeat_region | /rpt_family="At_rich" |
| repeat_region | 7192. .8658 |
| repeat_region | /gene="At2g43890" |
| repeat_region | /note="FE613.2" |
| repeat_region | join(7192. .7632,7743. .8032,8122. .8350,8440. .->8658) |
| repeat_region | /gene="At2g43890" |
| repeat_region | join(7192. .7632,7743. .8032,8122. .8350,8440. .8658) |
| repeat_region | /gene="At2g43890" |
| repeat_region | /codon_start=1 |
| repeat_region | /product="putative polygalacturonase" |
| repeat_region | /protein_id="AAC23398.1" |
| repeat_region | /db_xref="GI:3312847" |
| repeat_region | /translation="MDNKKLVAVILFFSESLMKSTPAASNVYVSGAKPDGRDTS |
| repeat_region | TKFLDAMQACRASAAYVTVPGRSTFLKPEFRPCRSRTTPIYGTIVAPSDYNG |
| repeat_region | LGSSGWSLLEFVAVNRSLIIIGGTLIDARGASWACRKGSKSPVGASMTFMANADVNS |
| repeat_region | GTLINSQTTFLVNSCNVNIIVRVKLVAPDQSGIDLVQSGAVTVDTGTFHTD |
| repeat_region | DCGISLPGRRNLYMSKLKCGPHGISISGLRDNAGVENTLLINSVFSGDNGVAI |
| repeat_region | KTNARQSTGEVNVFLFQNLIRKNYONPIIVDQNYCPNQGCKROSGVAKISQVYRNII |
| repeat_region | QGSIRYQALITFDCSRNMQCAIRLHDKIKTLFNGRSATSTCKNLINQVAGVYMQGL |
| repeat_region | " |
| repeat_region | complement(9101. .14090) |

| gene | complement(9101, .14090) | /gene="AT2943900" | /note="FEI3.3" | complement(join(<9101, .9952,10049, .10409,10500, .10523,10717, .10964,11093, .11133,11219, .11384,11587, .11671,11768, .11895,12004, .12811,12910, .13162,13359, .>14990)) |
|------|--|-------------------|--|---|
| mRNA | complement(join(9101, .9952,10049, .10409,10500, .10523,10717, .10964,11093, .11133,11219, .11384,11587, .11671,11768, .11895,12004, .12811,12910, .13162,13359, .>14990)) | /gene="AT2943900" | complement(join(9101, .9952,10049, .10409,10500, .10523,10717, .10964,11093, .11133,11219, .11384,11587, .11671,11768, .11895,12004, .12811,12910, .13162,13359, .>14990)) | |
| CDS | complement(join(9101, .9952,10049, .10409,10500, .10523,10717, .10964,11093, .11133,11219, .11384,11587, .11671,11768, .11895,12004, .12811,12910, .13162,13359, .>14990)) | /gene="AT2943900" | complement(join(9101, .9952,10049, .10409,10500, .10523,10717, .10964,11093, .11133,11219, .11384,11587, .11671,11768, .11895,12004, .12811,12910, .13162,13359, .>14990)) | |

| repeat_region | 14103..14213 |
|---------------|---|
| gene | /rpt_family="(GAGAA)n" complement(14748..17292) /gene="ATG243910" /note="FBE13.4; supported by full length cDNA: Ceres:24370" |
| mRNA | complement(Join<14748..14963,15285..15363,15476..15522) |

```

mRNA complement(join(<14748..14963,15285..15363,15476..15522
15617..15716,15874..15974,16059..16145,16677..16807,
17119..>17292))

```

GDDCIALSQGSRNWIERVNCGPGHGISIGSLGDYANEQVQNTVTSSVFTKTQNGV


```

Query Match      21.8%: Score 121.8; DB 8; Length 109741:
Best Local Similarity 79.6%: Pred. No. 1.4e-20:
Matches 144; Conservative 0; Mismatches 37; Indels 0; Gaps 0:

QY 149 AACAAAGATACCACCTGCATTTATGCGTCTTTAGAGCCTATCAAGCAATGAAAGAAC 208
    ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db 93752 AATGTAGATACCAGTGTGATGAGTGGCTGCGCTTGAGACCTATCAAGATTAACGAAAGAAC 93811
QY 209 TGTCAAGGCTTATGGAATTCACCTGGGAACCGAAGTGTGCAAAAAGATTTTAGCTTCATG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93812 TGTGAAACCTATGATGATTCACCTTGGAACAGAAATGTGA AAAAATGTTGGCTCATG 93871
QY 269 AATTAAAGCATTCGATCTTTATACACTAAATATGAGAAATCTGCATTGGCACTACTAAT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93872 ACTCAAGTCTCTTATCTCTACACTGATGAACATGAGAAATCTGCTTGCAATATTGAT 93931
QY 329 G 329
Db 93932 G 93932

RESULT 13
SPAC56F8 LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S.pombe)
Protein coding regions (CDS) have been predicted with the help of
computer analysis using the genefinder program in Pombase (an ACED
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC5H10.01c. SP (S
pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid c56F8 is overlapped at the 5' end by cosmid c30D11 and at
the 3' end by cosmid c22A12.
Location/Qualifiers
1..37093
/organism="Schizosaccharomyces pombe"
/strain="972h-"
/db_xref="taxon:4896"
/chromosome="I"
/contig="I1"

```

```
gene
/cldone="cosmid c56f8"
1. 1929
/gene="SPAC56f8.01"
/note="SPAC30D11.01c"
CDS
1. 1929
/partial
/gene="SPAC56f8.01"
/note="SPAC56f8.01, len: 642, SIMILARITY:Aspergillus
niger, AGLU_ASPNG, alpha glucosidase precursor, (1428 aa),
fasta scores: opt: 1586, E():0, (42.5% identity in 650
aa)"
/codon_start=1
/label="SPAC56f8.01
/product="putative family 31 glycosyl hydrolase (alpha
glucosidase) precursor"
/protein_id="CAA93572.1"
/db_xref="GI:1204223"
/db_xref="SWISS-PROT:Q09901"
/translation="DLMRSYLNASNIPREGWNDIDYMSSEPRTPVNSAPPNOT
LDFFRSLDESHQVVPYLDPAIVANPNKSKADRYVYVYSGFEDNIFIKPNQSAVYG
MAMGFVYVPTPTNPVLYQYKQGIILMSTARFSNYSIDLPFSGICLDMEPFSCIG
SCGSDILKLNVPFSLPGVDNRYSTYDEFNATVTEKSYSRASQSYKATAS
EKSHETPSESLINGKEPESLNPYPALDDTDTHDAQGVSPNATMHGNTLYNLF
NTYGESKISIFEALNSIQPNIRPELISRSTFVSGRYAAHMLGDNKSQMDLWYSIS
SILFENLGIPIWAGVGYNGNTDEELCARMAALGAPLPEYRNHNSLISPOEPFM
ASVAESRSATEIRVYSLPYWTLMHVASVDGPPMVARPLPEEPKOTSLASVDKQEMI
GTALLSPALEPNTTYIGGIPGNDITWDTNHSVINHDYDENTMSAPLGVNA
VRGNITPLQDPGTTTESRNPPSLIAMDNNCFASGSLYIDGISMQNTSLSVL
NSNTVTTCVSGTWSSPSLIANTITIGLSNPNTILFNGQQLSDYOYSDQJLSLNL
LDLTVDGAFSKNMTVWS"
1. 101
/gene="SPAC56f8.01"
/note="nominal overlap with cosmid SPAC30D11, EM:267961 S.
pombe chromosome 1"
4. 1812
/gene="SPAC56f8.01"
/note="Match to PF01055 Glyco_hydro_31, Glycosyl
hydrolases family 31 Score 977.93"
2567. 7120
/gene="SPAC56f8.02"
2567. 7120
/gene="SPAC56f8.02"
/note="SPAC56f8.02"
SIMILARITY:Schizosaccharomyces pombe, YAB4, SCHPO,
hypothenical 162.4 kd protein c22f3.04 in chromosome 1.,
(1428 aa), fasta scores: opt: 1491, E():0, (30.0% identity
in 1372 aa)"
/label="SPAC56f8.02
/codon_start=1
/product="hypothenical integral membrane protein"
/protein_id="CAA93573.1"
/db_xref="GI:1204224"
/db_xref="SWISS-PROT:Q10250"
/translation="MNOPNPGNGVGYKYPVQSGIPANSEATNFQONNSRENKSEC
ELRONSIAASGAYPNGMYAGAHNNHNNHNTYMGDHMEEYVGDLDVEPIAS
SDIYATPDGKDFGSLHVDOTLDQOMAHFAKOOHPLEPPIPPVTDLSKITEMO
FTNIAAVLYRYGVSAKKTAFIITIDNGKEPTSTWKLASRAKVAOVYRDKGLR
SDRYVIMRDCBALIDVYSLFGCTIAGVAVPIINRPDIYNELSLITTSKRALITD
ANLAFKQDILNAKLIHMPKNVEMKNTNEFGFHLKKAEMPPLOVPDATYERSPL
GELHGVNISHRTILHQNCLAIHATAVAYSDKLIVLSIDRETEGSLKGLITL
DLRQAIIGLIVLTVESGYTWCPONAVYVPIAMNALTARYESFMLDVAALKTI
AYNQNDCRKATIGFSKSHVDLSLRMCMVCLNVDFEOTIVSDRMLKPLGNPNPA
TFVPLIDPEHGAWISMKDWIGEEMSPKGFSPRPNENISEVILEKALKINY
VYLADDDARQSKHPNTIRYGAWYPPVDTALIVPEYOVLPLNIVGSIWDS
LSGFFALPKOTEIHFARISFISDTQPIPSQOEILRTGLGFIRKGVYELGTE
DLQOKVEMVNGKODITFFHHYSHLVNIMRVSKFLTDCSADIPVNSHLPVL
ESPANITPEANGQVYINGLDILITECECECLEHQAIVYCVILCAPTLPRVK
NGRQIENMCCRABEHGTLPELVYKFAVERAVNLVGEDAIDGIMSSYSGIRONL
LSDOLOYSGTIDRSLRYDAKTSYDISCHTMLQILOIRAKNMDIATYIDRSGE
GNITWKRKFDQVATITIRLOKKYIRGRVYVLMYTISEPVYALACPTIGIPI
VPLDHRLESDVPAFLIKHYVAVLVASEDDTALRAATTSQHLKQSMAKVAVL
PSFVTSKISKQTSIKELNVKLDPICDPAFPALVMAFMSPDHRLTLTVAHDTLLS
ICQIHETCOMTHKRPLILGHVSRMSGIGFPTCLMGVFLGTTVILSPVDRANPDLIL
```

```
gene
/gene="SPAC56f8.03"
7948. 11187
/gene="SPAC56f8.03"
/note="SPAC56f8.03, len:1079, SIMILARITY:Saccharomyces
cerevisiae, IF2P_YEAST, translation initiation factor
if-2, (1002 aa), fasta scores: opt: 3165, E():0, (52.8%
identity in 1015 aa)"
/codon_start=1
/label="SPAC56f8.03
/product="translation initiation factor if-2"
/protein_id="CAA93574.1"
/db_xref="GI:1204225"
/db_xref="SWISS-PROT:Q10251"
/translation="MGKRGKSGVADWEDDLDGEDISGONEYIDNTSQSPONDELAEK
SEMLAVSSKRTSKKRRKNGKNKNOVSDDESQELSPQGRKELTAVELDDEFDY
KPKRKGKRSKVEEDDEPOEIESPOGPELTAVELDDEFDYKPKKGGKKAQN
NNSEAAAPPEIPEYRVKTKKEERKKRRELRKKQOAKKSGTGEDYKLAASSVS
EVIDSTPAMESAKGKQAGKRRKGPNTYALOKMLEEKAREBEORIREEARLAE
EERLAPVEARKEEARLKKKEERKKREEMKAOCKYLSKQKQDQALQORLEQAL
SGRVAGLSNGEKQKQPVYTNKKKSRSSTSSISSGIIESSPAISIVDEQKQSKD
DSKREKEREKEENEAEAEVFDWDEADALEPVAENNEVVEKRTDILKSQVE
HSTKDEKSTKDVDDIPQAAPEASVNSVSDLEPCCILGHVDTGKRLDNLNRSN
VQGEAGGITQOIGATYVPIESIKQTKYVNNKGIQVNIPIGLIIDPGHISFINLR
SRGTSCLNIAIVYIDIMHGLEPOTIESILLDOKTPPEVANKRYDRAVMSIDNA
IODLSKOKKAIORFERDRESIIQINQNGIAMAALYENKMLGRVSYLSPVSAQGE
GVDPVALLISUTQPSMRIVITITLBECTVEVYIEGATIDVYISNGVLSGDR
IVCGMGGPIITTVRALTPQPKEMKRVASAVVHREITAKANGYTCANDLEKAVAGS
RLVAVGPDDDEEDLAEIMEDLENLIGRIDTSGIGSVASTLSLEALEFLKQMKI
PVASVIGIYVKKDVARCATMLEKAEVAMLCFDVYKDRADAEQAGVIFESANV
IYHLPFAFHAOKRIEOKREESDVAVPVCYLTVAAKNKADPITLIGDVYEGVARI
NPIVAVKOLPNEGPOITELGRVASELNNHKKVPDKYKQAGVAMKRESSGQILF
GRQVTSDALYSHITRQSDLSLKDPAFRQEVSRDEMOQLIQKLKLGIT"
8468. 9168
/gene="SPAC56f8.03"
/note="Match to PF00992 Troponin, Troponin Score 22.61"
9418. 9441
/gene="SPAC56f8.03"
/note="PS00017 ATP/GTP-binding site motif A"
complement(11626..12702)
/gene="SPAC56f8.04c"
complement(11626..12702)
/gene="SPAC56f8.04c"
/note="SPAC56f8.04c, len:358, SIMILARITY:Saccharomyces
cerevisiae, COQ2_YEAST,
para-hydroxybenzoate--polyprenyltransferase, mitochondrial
precursor(ec 2.5.1.-), (372 aa), fasta scores: opt: 909,
E():0, (48.3% identity in 296 aa)"
/codon_start=1
/label="SPAC56f8.04c
/product="para-hydroxybenzoate--polyprenyltransferase,
mitochondrial precursor(ec 2.5.1.-)"
/protein_id="CAA93575.1"
/db_xref="GI:1204226"
/db_xref="SWISS-PROT:Q10252"
/translation="MIKIPASARFELRPPMSAVAIFOAVKIKPLQJLRTSSSVY
PNLISPSKSKMDPLSKRMQYVETISRASPGCTYLLSPCWMISLMAVAVDSSLVN
VTLMSLGLIGLQMPYIKLGVALVPCVILYPAKRTTYIPQVVLGTFQYGVAKMPA
LAEACQENNSVAVPLVSTISWLVLDITTYAHODKSDVKANITYSTALFQGNTPVL
CGAALQIOLATLAGIMNGGQPVETVIGVAAAYRILSSMIVYDLPDPCDFRMEKRN
SNNGYVAAAIADMLAKSFYDS"
complement(11653..12411)
/gene="SPAC56f8.04c"
/note="Match to PF01040 COX10_ctab_cyoE, Cytochrome c
oxidase assembly factor Score 444.71"
```

misc_feature

gene complement(join(13097..13461,13595..13802,13917..14084,
14129..14176,14332..14370,14511..14570))
/gene="SPAC56F8.05c"
CDS complement(join(13097..13461,13595..13802,13917..14084,
Matches 290; Conservative 0; Mismatches 246; Indels 15; Gaps 1;
Query Match 21.0%; Score 117.4; DB 8; Length 37093;
Best Local Similarity 52.6%; Pred. NO. 2.1e-19;
Matches 290; Conservative 0; Mismatches 246; Indels 15; Gaps 1;
OY 1 TGATGTCCACCAATTTATTTATGATACGATATATTCCTCAATTTGTGAACGACTGTC 60
DB 22053 TTATGCTTACTCAATTTTATGATGATAGACATTTTATTCGATGGCTGATAAAGTGC 22112
OY 61 GCCAAATTGGAAATTAACGTCCTATTTGCTATTTGCTGGAATTTATCCCATTAATTAATTAACAAG 120
DB 22113 GTGACAGAGGAATTAATATCCCTATATTTCCGGGCATTAATGCTATTTACAGGCTGGGATT 22172
OY 121 GCTTTATCCGATGACTGGGTTTTCGAAAACAAGATACGAGCTGACATTTATGGCTGCTT 180
DB 22173 CCTTATTCGAGAGCGAATGAGCGGCTGTAAATTTCCCGACATTTTATGATATCTC 22232
OY 181 TAGAGCCTATCAAGAACATGAGAAAGCTGTCAAGCTTTATGAAATTCACCTGGGAACTC 240
DB 22233 TAGTCCAGTTTAAAGAGATGATGAGAGAGTCCGTAGCGGTGTGAGCTCATATGCTG 22292
OY 241 AAATGTGCAAAAAGATTTAGTCTCATGATGAATTAAGACATTTGCTTTATACATAAATA 300
DB 22293 AAATGTGCGTAAGCTTATGCTAGTGGCATTAAGAGACTTCAATTTTACACATATGAATT 22352
OY 301 TCGAGAAATGTGCTATTTGGCAATCAATGAACCTTGCCCTAATTTGAAGAGTCCAAAGTTT 360
DB 22353 TAAAAAAGGCGTTTAAATATTTATGAAAGATAGATTTATGATGAATAAATCTGGCTC 22412
OY 361 CTAGGTCTTACCTTGGAGAGCGCCCTGGCAA-----ATGTTTCCGCTGTTA 405
DB 22413 CTATAGTGGTACTAATTAAGCTGAGTTTACCAATGCTTCCAGTCAAGATCGTGGATTA 22472
OY 406 AAGAAGATGTCCGTCCTCAATCTTTTGGGCAATGACCAAAAAGCTCATATCAAGAGCA 465
DB 22473 ATGAAGGTGTAGCGCCATTTCTGGGCGACCTGTANTGAAGTTATGCTCGCGTACG 22532
OY 466 TAGATGGGATCAATACCCCATGSGCGTTGGGGTATTCCTGTAAATCATCATATGATG 525
DB 22533 ATCAGTGGGAGCAATTAACCGCATGTGTTGGGTGACTCTGTACCCCTGCTTTTGGCG 22592
OY 526 CATTATCTGAT 536
DB 22593 AATTGATGCT 22603
RESULT 14
AX069359 1971 bp DNA linear PAT 25-JAN-2001
LOCUS AX069359
DEFINITION Sequence 23 from Patent WO0102600.
ACCESSION AX069355
VERSION AX069359.1 GI:12579222
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1971)
AUTHORS Yuan,C.S.
TITLE Detection of analytes using attenuated enzymes
JOURNAL Patent: WO 0102600-A 23 11-JAN-2001;
GENERAL NOTES location/Qualifiers
FEATURES
source 1..1971
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Human methylentetrahydrofolate reductase (MTHFR)
gene; exons 1-8"
BASE COUNT 450 a 566 c 570 g 385 t

ORIGIN
Query Match 20.9%; Score 116.4; DB 6; Length 1971;
Best Local Similarity 53.3%; Pred. NO. 4.7e-19;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
OY 1 TGAATGTCCACCAATTTATTTATGATACGATATATTCCTCAATTTGTGAACGACTGTC 60
DB 671 TCATCATCCGACGAGCTTTTCTTTGAGCTGACACATTTCTCCGCTTTGTGAAGCATCA 730
OY 61 GCCAAATTGGAAATTAACGTCCTATTTGCTATTTGCTGGAATTTATCCCATTAATTAACAAG 120
DB 731 CCGACATGGGCAATCACTTGCCCATTCGTCGCCGATCTTCCATTCACAGGCTACACCT 790
OY 121 GCTTTATCCGATGACTGGGTTTTCGAAAACAAGATACGAGCTGACATTTATGGCTGCTT 180
DB 791 CCTTCGGGAGCTTGTGAAGCTGTCCAAAGCTGAGAGTCCACAGAGATCAAGAGCTGGA 850
OY 181 TAGAGCCTATCAAGACATGAGAAAGCTGTCAAGCTTTATGAAATTCACCTGGGAACCTG 240
DB 851 TTGAGCCATCAAGACATGAGAAAGCTGTCCATTCGCACTATGCAATGAGAGCTGGCGTGA 910
OY 241 AAATGTGCAAAAAGATTTAGCTCATGATGAATTAAGACAT---TGCATCTTTATACATAA 297
DB 911 GCCTGTCCAGAGAGCTTCTGGCCAGTGGCTGTGTCAGGCTCCACATTTACACCTCA 970
OY 298 ATATGGAAGATCTGATTTGGCAATTAATGAACCTTGCCCTAATTTGAAGAGTCCAAAG 357
DB 971 ACCGGGAATGGCTTACACAGAGGTGTGAAGCGCTTG6GATGTGACTGAGACCCCA 1030
OY 358 TTTTAGGTCTTCTTACCTTGGAGAGCGCCCTGCAAAATGTTTCCGTTTAAAGATGTC 417
DB 1031 GGGCTC---CCCTACCTCGGGCTCTCATGCTCCACCCCAAGCGCCGAGAGAAATGTAC 1087
OY 418 GTCAATCTTTTGGGCAATGACCAAAAAGCTCATATTAAGAGCAATAGATGGATC 477
DB 1088 GTCCCATCTTCTGGGCTCCAGACCAAGAGTTATCATCTACCGTACAGAGAGGGAGC 1147
OY 478 AATACCCACATGGGGTGGGTGATTCCTGTATTCATCATATGATGATATATCTAT 537
DB 1148 AGTTCCTTACGCGCGCTGGGCAATTCCTCTCCCTGCTTTGGGAGCTGAAGACT 1207
OY 538 ATCACTTCAAT 547
DB 1208 ACTACTCTT 1217
RESULT 15
ARI44956 2187 bp DNA linear PAT 08-AUG-2001
LOCUS ARI44956
DEFINITION Sequence 1 from patent US 6210950.
ACCESSION ARI44956
VERSION ARI44956.1 GI:15106823
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2187)
AUTHORS Johnson,W.G. and Stenroos,E.Scott.
TITLE Methods for diagnosing, preventing, and treating developmental
disorders due to a combination of genetic and environmental factors
JOURNAL Patent: US 6210950-A 1 03-APR-2001;
GENERAL NOTES location/Qualifiers
FEATURES
source 1..2187
/organism="unknown"
BASE COUNT 479 a 655 c 616 g 437 t
ORIGIN
Query Match 20.9%; Score 116.4; DB 6; Length 2187;
Best Local Similarity 53.3%; Pred. NO. 4.7e-19;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
OY 1 TGAATGTCCACCAATTTATTTATGATACGATATATTCCTCAATTTGTGAACGACTGTC 60

```
Db      674 TCATCATCACGAGCTTTTCTTGGAGCTGACATCTTCGCCGTTTGTGAAGCATGCA 733
QY      61  GCCAATTGGAAATTAACGTGTCTTATGTACTGGAAATTATGCCATTATTAATTCAGG 120
Db      734  CCGACATGGGACATCACTTCCCATCTGCCATCTGCCGGATCTTCCATCCAGGGCTACCACT 793
QY      121  GCTTTATCCGATGACTGGGTTTGTGCAAAACAAAGATACCAAGCTGACATTATGGCTGCTT 180
Db      794  CCTTCGGGAGCTGTGGAAGCTGTCCAAAGCTGAGGTGCCACAGAGATCAAGAGACGTGA 853
QY      181  TAAGGCTTTCAGAGCAATGAAGCTGTCAAGGCTTATGAATTTCACCTGGGACTG 240
Db      854  TTAGGCCAATCAAGACATGCTGCCATCCGAACTATGGATCGAGCTGGCCGTGA 913
QY      241  AAATGTCAAAAAGATTTTATGCTCATGGAATTAGACAT--TGCACTTTATACACTAA 297
Db      914  GCGTGTGCCAGAGCTTGTGGCCAGTGGCTGTGGCCAGGCCCTCCACTCTACACCTCA 973
QY      298  ATATGAGAAATCTGCATTTGGCAATTAATGAACCTTGGGCTAATTGAAGAGTCCAAAG 357
Db      974  ACCGCGATGGCTACCAAGAGGTGTGAAGCGCTGGGGATGTGACTGAGAGACCCA 1033
QY      358  TTCTAGGTCTTACTCTGGAGAGCGCCTGCAAATGTTTCCGTGTTAAGAAGATGCC 417
Db      1034  GGCCTC--CCCTACCTTGGGCTCTCAGTGGCCACCCCAAGCGCGAGAGAGATGTAC 1090
QY      418  GTCCAACTTTTGGGCAATCGACCAAAAGCTACATATCAAGGACCATAGATGGATC 477
Db      1091  GTCCCAATCTTGTGGGCTTCCAGCAAAAGATTACATCTACCTACCCAGAGATGGAGC 1150
QY      478  AATACCCACATGGGCTTGGGCTGATTCCTGTAATCCATCATATGCTGATTTATCGATT 537
Db      1151  AGTTCCTTACGGCGCTGGGCAATCTCTTCCCTTGCTTGGGAGCTGAAGGACT 1210
QY      538  ATCAGTTCAT 547
Db      1211  ACTACTCTT 1220
```

Search completed: February 14, 2003, 21:20:23
Job time : 3017 secs

1000

1000

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 10:25:29 ; Search time 266 Seconds

(without alignments)
4724.117 Million cell updates/sec

Title: us-09-720-451-5

Perfect score: 558
Sequence: 1 tgaatgcacccaattattt.....tcagttcatgcgcacacgrr 558

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------------------|
| 1 | 558 | 100.0 | 558 | 21 | AA50061 Soybean 5,10-methyl |
| 2 | 378.8 | 67.9 | 1821 | 21 | AA43311 Arabidopsis thaliana |
| 3 | 378.8 | 67.9 | 1942 | 21 | AA44918 Arabidopsis thaliana |
| 4 | 358.4 | 64.2 | 1493 | 21 | AA44829 Arabidopsis thaliana |
| 5 | 346 | 62.0 | 1335 | 21 | AA35967 Arabidopsis thaliana |
| 6 | 139.6 | 25.0 | 451 | 21 | AA50062 Wheat 5,10-methyl |
| 7 | 116.4 | 20.9 | 2187 | 22 | AA50062 Human methyl |
| 8 | 116.4 | 20.9 | 2219 | 17 | AA709689 Human methyl |
| 9 | 116.4 | 20.9 | 2219 | 22 | AA20463 Human methyl |

| | | | | | |
|----|-------|------|--------|----|------------------------------|
| 10 | 116.4 | 20.9 | 2220 | 17 | AA709694 Human methyl |
| 11 | 116.4 | 20.9 | 2220 | 21 | AA50633 Human methyl |
| 12 | 116.4 | 20.9 | 2220 | 22 | AA20462 DNA encoding human |
| 13 | 116.4 | 20.9 | 2220 | 24 | ABK13501 Human methyl |
| 14 | 116.4 | 20.9 | 2220 | 24 | ABK13539 Human methyl |
| 15 | 116.4 | 20.9 | 2220 | 24 | ABK13540 Human methyl |
| 16 | 116.4 | 20.9 | 2220 | 24 | ABK13541 Human methyl |
| 17 | 116.4 | 20.9 | 2220 | 24 | ABK13542 Human methyl |
| 18 | 116.4 | 20.9 | 2220 | 24 | ABK13545 Human methyl |
| 19 | 116.4 | 20.9 | 2220 | 24 | ABK13546 Human methyl |
| 20 | 116.4 | 20.9 | 2220 | 24 | ABK13547 Human methyl |
| 21 | 116.4 | 20.9 | 2220 | 24 | ABK13550 Human methyl |
| 22 | 114.8 | 20.6 | 2220 | 24 | ABK13543 Human methyl |
| 23 | 114.8 | 20.6 | 2220 | 24 | ABK13544 Human methyl |
| 24 | 114.8 | 20.6 | 2220 | 24 | ABK13548 Human methyl |
| 25 | 97 | 17.4 | 159 | 24 | ABU75279 Corn tassels-derive |
| 26 | 53.4 | 9.6 | 640681 | 24 | ABA92787 Buchnera sp. genom |
| 27 | 51.8 | 9.3 | 235 | 22 | AA20491 Mouse MTHFR gene e |
| 28 | 51.8 | 9.3 | 235 | 24 | ABK13519 Mouse methyl |
| 29 | 51.8 | 9.3 | 3181 | 21 | AA50635 Mouse methyl |
| 30 | 48.6 | 8.7 | 235 | 22 | AA20480 Human MTHFR gene e |
| 31 | 48.6 | 8.7 | 235 | 24 | ABK13508 Human methyl |
| 32 | 48.6 | 8.7 | 3283 | 21 | AA50634 Human methyl |
| 33 | 47 | 8.4 | 23748 | 22 | AAU05223 Human reproductiv |
| 34 | 47 | 8.4 | 23748 | 23 | ABU98112 Human testicular a |
| 35 | 43.6 | 7.8 | 294 | 22 | AA20478 Human MTHFR gene e |
| 36 | 43.6 | 7.8 | 294 | 24 | ABK13506 Mouse methyl |
| 37 | 41.8 | 7.5 | 294 | 22 | AA20489 Mouse MTHFR gene e |
| 38 | 41.8 | 7.5 | 294 | 24 | ABK13517 Mouse methyl |
| 39 | 41.8 | 7.5 | 1082 | 23 | ABK13089 Drosophila melanog |
| 40 | 41.8 | 7.5 | 3213 | 23 | AAV52273 Strepococcus pneu |
| 41 | 40.4 | 7.2 | 28862 | 22 | AAV52273 Soybean 515002 reg |
| 42 | 40 | 7.2 | 127197 | 19 | AA161370 DNA encoding a S. |
| 43 | 38.8 | 7.0 | 1449 | 19 | AA198588 Human serine palmi |
| 44 | 37 | 6.6 | 646 | 24 | AA148406 Human serine palmi |
| 45 | 37 | 6.6 | 646 | 24 | AA234462 |

ALIGNMENTS

| | | |
|----------|--|---|
| RESULT 1 | AA50061 | standard; cDNA; 558 BP. |
| ID | AA50061 | |
| XX | AA50061 | |
| XX | AA50061 | |
| DT | 04-MAY-2000 | (first entry) |
| DE | Soybean 5,10-methyltetrahydrofolate reductase cDNA clone. | |
| XX | 5,10-methyltetrahydrofolate reductase; MTHFR; soybean; inhibitor; | |
| KW | methionine synthesis; tetrahydrofolate metabolism enzyme; herbicide; ss. | |
| OS | Glycine max. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | mat_peptide | 3..557 |
| FT | | /*tag= a |
| FT | | /partial |
| FT | /product= | "5,10-methyltetrahydrofolate reductase" |
| FT | /note= | "encodes amino acids Ile4,Arg188 of the enzyme" |
| FT | /EC_number= | "1.7.99.5" |
| PN | WO200004163-A1. | |
| XX | | |
| PD | 27-JAN-2000. | |
| XX | | |
| XX | | |
| PF | 14-JUL-1999; | 99MO-US15916. |
| XX | | |
| PR | 15-JUL-1998; | 98US-0092869. |
| XX | | |
| PA | (DUPO) DU PONT DE NEMOURS & CO E I. | |

XX Falco SC, Fomodu LO;
PI
XX
DR MPI: 2000-182429/16.
DR P-PSDB: AAY44740.
XX
XX Novel tetrahydrofolate metabolism enzyme used to alter the level of
PT tetrahydrofolate metabolism in plants and seeds -
XX
PS Claim 3: Page 32; 37pp: English.
XX
XX The present sequence is a cDNA clone sfl1_pk0017_d12 encoding
CC 5,10-methylenetetrahydrofolate reductase (MTHFR). The clone was
CC isolated from a sfl1 cDNA library which was prepared using soybean
CC immature flowers. MTHFR plays a role in the
CC synthesis of methionine. The present sequence is used in the
CC construction of a chimeric gene to alter the level of tetrahydrofolate
CC metabolism enzymes in plants. The enzyme may provide target to facilitate
CC design and/or identification of inhibitors that may be useful as
CC herbicides. The polynucleotide is also useful as a source of probes for
CC genetically and physically mapping the genes and as markers for traits
CC linked to the genes.
XX
SQ Sequence 558 BP; 170 A; 109 C; 116 G; 163 T; 0 other;
Query Match 100.0%; Score 558; DB 21; Length 558;
Best Local Similarity 100.0%; Pred. No. 9.1e-156;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGATGTCACCAATATTTATGATACGATATATCTCTCAAAATTTGACAGCTGTC 60
DB 1 TGATGTCACCAATATTTATGATACGATATATCTCTCAAAATTTGACAGCTGTC 60
QY 61 GCCAAATGGATAGCTGCTATTTGATACGGAATTAAGCCATTATATACAGG 120
DB 61 GCCAAATGGATAGCTGCTATTTGATACGGAATTAAGCCATTATATACAGG 120
QY 121 GCTTATCCGAGTACTGSGTCTTGGCAAAACAAGATACAGCTGATATGCTGCTT 180
DB 121 GCTTATCCGAGTACTGSGTCTTGGCAAAACAAGATACAGCTGATATGCTGCTT 180
QY 181 TAGAGCCATCAAGCAATGAGAGCTGTACAGCTTATGAAATTCACCTGGAGCTG 240
DB 181 TAGAGCCATCAAGCAATGAGAGCTGTACAGCTTATGAAATTCACCTGGAGCTG 240
QY 241 AAATGTGCAAAAGATTTTACCTCATGAGAAATTAAGCAATTCATTTATACATAATA 300
DB 241 AAATGTGCAAAAGATTTTACCTCATGAGAAATTAAGCAATTCATTTATACATAATA 300
QY 301 TGGAGAAATCTGATTTGCAATGCAATGACCTGGCCTATTTGAAAGAGTCAAAAGTTT 360
DB 301 TGGAGAAATCTGATTTGCAATGCAATGACCTGGCCTATTTGAAAGAGTCAAAAGTTT 360
QY 361 CTAGGTCTTACCTTGGAGAGCGCCGCAATGTTTCCGCTTAAGAAATATCCCTC 420
DB 361 CTAGGTCTTACCTTGGAGAGCGCCGCAATGTTTCCGCTTAAGAAATATCCCTC 420
QY 421 CAATCTTTTGGCAAAATGACAAAAAGCTACATATCAAGACCAATAGATGGATCAAT 480
DB 421 CAATCTTTTGGCAAAATGACAAAAAGCTACATATCAAGACCAATAGATGGATCAAT 480
QY 481 ACCGCAATGGGCTTGGGGGTGATCTGTAATCCCAATGAGGCAATATTCATATATC 540
DB 481 ACCGCAATGGGCTTGGGGGTGATCTGTAATCCCAATGAGGCAATATTCATATATC 540
QY 541 AGTTCATGCGGCCACGTG 558
DB 541 AGTTCATGCGGCCACGTG 558
RESULT 2
AAC43311
ID AAC43311 standard; DNA: 1821 BP.

XX AAC43311;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38803.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
PD
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126268.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131445.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140821.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144003.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145915.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145937.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 67.9%; Score 378.8; DB 21; Length 1821;
Best local Similarity 79.9%; Pred. No. 2,5e-102;
Matches 446; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1 TGAATGTCACCCATATTTATATGATACGATATATCTCTCAAAATTTGACACAGCTG 60
DB 602 TGAATGTTACTCAGCTTTCTATGATAGTATATATCTCAAGTTTGATGATATTTG 661
QY 61 GCCAAATTTGAAATACGATGTCCTATTTGACCTGGAATTAATGATTAATATTCAGG 120
DB 662 GGCAAATTTGGAATAGTAGTGTCCCATTTGCTCGAATTAATGATTAATATTCAGAG 721
QY 121 GCATTATCCGATGACGTGGTTTTCACAAACAAAGATACAGCTACATTAATGCTGCTT 180
DB 722 GCATTATCCGATGACGTGGTTTTCACAAACAAAGATAGATTAATGATTAATGCTGCTT 781
QY 181 TAAAGCCTATCAAGACATTAAGACGCTTCAAGCCTTATGAAATTTACTGAGACG 240
DB 782 TGAAGCCTATCAAGACATTAAGACGCTTGAAGACCTATGATTAATCACTTGAGACG 841

```
OY 241 AAATGTCAAAAGATTAGTCATGGAATTAGACATTGCATCTTATTACACTAATA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 842 AGATGTGTAAAGATTTGGCTCATGAGTCAAGTCTCTCATCTACATGAGACA 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 301 TGSAGAAATCTGCATTGAGCAATACATGACCTTGGCCTTAATGGAAGTCCAAAGTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 902 TGSAGAAATCTGCCTTCATGATGAAATCTTGGTATGATGATGATGATGCCAAATTT 961
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 361 CTAGTCTTACCTTGAGAGAGCCCTGCAAGATGTTTCCTGCTTAAGAAGATGCCCTC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 962 CTCGTTCTTTACCTTGAGAGAGCCCTGCAAGATGTTTCCTGCTTAAGAAGATGCCCTC 1021
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 421 CAATCTTTTGGCAAAATCGAACCAAAAGCTACATATCAGAGACCATTAGATGGATCAAT 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1022 CCAATTTCTGGGCAAAACCGTCCAAAGAGCTACATTTCTAGAACCAAGGCGTGGAGATT 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 481 ACCCAATGAGCGCTTGGGCTGATTCCTGTAATCCATCATATGATGATTCATTCATTC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1082 TCCCAAGAGCCGCTGGGGTGATTCACGCACTGCTTATATGATGATGATGATGATC 1141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 541 AGTTCATGCGGCGCACGCTG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1142 AGTTCATGCTCGCGAG 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AAC49168
ID AAC49168 standard; DNA; 1942 BP.
XX
AC AAC49168;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SPQ ID NO: 60181.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130440.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
```


PD 06-SEP-2000.
XX
PF 25-FEB-2000. 200DEP-0301439.
XX
PR 25-FEB-1999. 990S-0121825.
PR 05-MAR-1999. 990S-0123160.
PR 09-MAR-1999. 990S-0123548.
PR 23-MAR-1999. 990S-0125788.
PR 25-MAR-1999. 990S-0126254.
PR 29-MAR-1999. 990S-0126725.
PR 01-APR-1999. 990S-0127462.
PR 06-APR-1999. 990S-0128224.
PR 08-APR-1999. 990S-0128714.
PR 16-APR-1999. 990S-0129845.
PR 19-APR-1999. 990S-0130077.
PR 21-APR-1999. 990S-0130510.
PR 23-APR-1999. 990S-0130449.
PR 28-APR-1999. 990S-0130891.
PR 30-APR-1999. 990S-0131449.
PR 04-MAY-1999. 990S-0132407.
PR 05-MAY-1999. 990S-0132485.
PR 06-MAY-1999. 990S-0132486.
PR 07-MAY-1999. 990S-0132863.
PR 11-MAY-1999. 990S-0134256.
PR 14-MAY-1999. 990S-0134218.
PR 14-MAY-1999. 990S-0134231.
PR 14-MAY-1999. 990S-0134231.
PR 18-MAY-1999. 990S-0134370.
PR 19-MAY-1999. 990S-0134768.
PR 20-MAY-1999. 990S-0134941.
PR 21-MAY-1999. 990S-0135124.
PR 24-MAY-1999. 990S-0135353.
PR 25-MAY-1999. 990S-0136021.
PR 27-MAY-1999. 990S-0136382.
PR 28-MAY-1999. 990S-0136722.
PR 01-JUN-1999. 990S-0137242.
PR 03-JUN-1999. 990S-0137358.
PR 04-JUN-1999. 990S-0137502.
PR 07-JUN-1999. 990S-0137734.
PR 08-JUN-1999. 990S-0138094.
PR 10-JUN-1999. 990S-0138540.
PR 14-JUN-1999. 990S-0138847.
PR 16-JUN-1999. 990S-0139432.
PR 17-JUN-1999. 990S-0139433.
PR 18-JUN-1999. 990S-0139442.
PR 18-JUN-1999. 990S-0139454.
PR 18-JUN-1999. 990S-0139455.
PR 18-JUN-1999. 990S-0139456.
PR 18-JUN-1999. 990S-0139457.
PR 18-JUN-1999. 990S-0139458.
PR 18-JUN-1999. 990S-0139459.
PR 18-JUN-1999. 990S-0139460.
PR 18-JUN-1999. 990S-0139461.
PR 18-JUN-1999. 990S-0139462.
PR 18-JUN-1999. 990S-0139463.
PR 18-JUN-1999. 990S-0139750.
PR 21-JUN-1999. 990S-0139817.
PR 22-JUN-1999. 990S-0139817.
PR 23-JUN-1999. 990S-0139889.
PR 23-JUN-1999. 990S-0140353.
PR 24-JUN-1999. 990S-0140354.
PR 28-JUN-1999. 990S-0140655.
PR 29-JUN-1999. 990S-0140923.
PR 30-JUN-1999. 990S-0141287.
PR 01-JUL-1999. 990S-0141287.
PR 01-JUL-1999. 990S-0141842.
PR 02-JUL-1999. 990S-0142154.
PR 02-JUL-1999. 990S-0142055.

PR 06-JUL-1999. 990S-0142390.
PR 08-JUL-1999. 990S-0142803.
PR 09-JUL-1999. 990S-0142920.
PR 12-JUL-1999. 990S-0142977.
PR 13-JUL-1999. 990S-0143247.
PR 14-JUL-1999. 990S-0143624.
PR 15-JUL-1999. 990S-0144008.
PR 16-JUL-1999. 990S-0144082.
PR 16-JUL-1999. 990S-0144086.
PR 19-JUL-1999. 990S-0144325.
PR 19-JUL-1999. 990S-0144331.
PR 19-JUL-1999. 990S-0144332.
PR 19-JUL-1999. 990S-0144333.
PR 19-JUL-1999. 990S-0144334.
PR 19-JUL-1999. 990S-0144335.
PR 20-JUL-1999. 990S-0144335.
PR 20-JUL-1999. 990S-0144632.
PR 20-JUL-1999. 990S-0144632.
PR 21-JUL-1999. 990S-0144814.
PR 21-JUL-1999. 990S-0145086.
PR 22-JUL-1999. 990S-0145087.
PR 22-JUL-1999. 990S-0145087.
PR 22-JUL-1999. 990S-0145089.
PR 22-JUL-1999. 990S-0145192.
PR 23-JUL-1999. 990S-0145145.
PR 23-JUL-1999. 990S-0145218.
PR 26-JUL-1999. 990S-0145224.
PR 27-JUL-1999. 990S-0145276.
PR 27-JUL-1999. 990S-0145911.
PR 27-JUL-1999. 990S-0145918.
PR 28-JUL-1999. 990S-0145951.
PR 02-AUG-1999. 990S-0146368.
PR 02-AUG-1999. 990S-0146388.
PR 03-AUG-1999. 990S-0146388.
PR 03-AUG-1999. 990S-0147039.
PR 04-AUG-1999. 990S-0147204.
PR 05-AUG-1999. 990S-0147120.
PR 05-AUG-1999. 990S-0147120.
PR 06-AUG-1999. 990S-0147260.
PR 06-AUG-1999. 990S-0147303.
PR 06-AUG-1999. 990S-0147416.
PR 09-AUG-1999. 990S-0147935.
PR 09-AUG-1999. 990S-0147935.
PR 10-AUG-1999. 990S-0148171.
PR 11-AUG-1999. 990S-0148319.
PR 12-AUG-1999. 990S-0148341.
PR 13-AUG-1999. 990S-0148565.
PR 16-AUG-1999. 990S-0148684.
PR 17-AUG-1999. 990S-0149368.
PR 18-AUG-1999. 990S-0149175.
PR 20-AUG-1999. 990S-0149426.
PR 20-AUG-1999. 990S-0149723.
PR 20-AUG-1999. 990S-0149723.
PR 20-AUG-1999. 990S-0149723.
PR 23-AUG-1999. 990S-0149902.
PR 23-AUG-1999. 990S-0149902.
PR 25-AUG-1999. 990S-0150566.
PR 26-AUG-1999. 990S-0150884.
PR 27-AUG-1999. 990S-0151065.
PR 27-AUG-1999. 990S-0151066.
PR 30-AUG-1999. 990S-0151080.
PR 31-AUG-1999. 990S-0151303.
PR 01-SEP-1999. 990S-0151438.
PR 07-SEP-1999. 990S-0151539.
PR 10-SEP-1999. 990S-0153363.
PR 10-SEP-1999. 990S-0153079.
PR 13-SEP-1999. 990S-0153758.
PR 15-SEP-1999. 990S-0154018.
PR 16-SEP-1999. 990S-0154039.
PR 20-SEP-1999. 990S-0154779.
PR 22-SEP-1999. 990S-0155139.

| | | |
|----|--------------|---------------|
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158363. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 14-OCT-1999; | 99US-0159638. |
| PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160778. |
| PR | 21-OCT-1999; | 99US-0160814. |
| PR | 21-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. |
| PR | 22-OCT-1999; | 99US-0160989. |
| PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161406. |
| PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161520. |
| PR | 28-OCT-1999; | 99US-0161592. |
| PR | 28-OCT-1999; | 99US-0161593. |
| PR | 29-OCT-1999; | 99US-0162142. |

| | Query Match Similarity | 64.28; | Score 358.4; | DB 21; | Length 1493; |
|----|------------------------|--|--------------------|-----------------|------------------|
| | Best Local Similarity | 77.88; | Pred. No. 2,76-96; | | |
| | Matches 421; | Conservative | 1; | Mismatches 122; | Indels 0; Gaps 0 |
| QY | 1 | TGATTTGTCACCCAAATATATTTTATGATACAGGATATATTCCTCAATATTTTGGAAAGCTCTC | 60 | | |
| Db | 649 | TGATTTGTCACAGCTTTTCTATGATACAGATATATTCCTCAATATTTTGGAAATGATTTTC | 708 | | |
| QY | 61 | GCCAAATTTGGANTTAACGTTGCTATTTGTACTGTGATATTTAGCCCATTAATTAATACAAAG | 120 | | |
| Db | 709 | GCGAAATTCGGATTTATTTATTTGCCATTTGTTCTCGGAATTAATGCCATTTCCAACTTCAAGG | 768 | | |
| QY | 121 | GCTTTATCCGCAATGCATGCTGGTTTGTGGAAAACAAATATACAGCTGACATTAATGCTGCTT | 180 | | |
| Db | 769 | GGTTTCTTGCGTATTTGCTGGTTTGTGTAACACCAATATACCCGCTGAGCTACCTGGCTGCTT | 828 | | |
| QY | 181 | TAGAGCTTTCACAGCAGATATGAAGAAGCTGTACAGCTTATGAGATTTACACTCGGGAATGTG | 240 | | |
| Db | 829 | TAGAGCTTTCATTAAGATATATACAGAGCTGTTAAGCTTATGAGATTTACCTTTGGCAATAG | 888 | | |
| QY | 241 | AAATTTGGCAAAAGTTTTTAGCTCATGAGAAATTAACACTTCATCTTTATACCTTAATA | 300 | | |
| Db | 889 | AAATTTGGCAAAAGTTTTTGGCCCATGAGATCACTCCCTCATCTCTACACTTTGGAAGC | 948 | | |
| QY | 301 | TGGAGAAATCTGCATTTGSCAATTAATGACCTTGGCTTAATTTGAAGAGTCCAAAGTTT | 360 | | |
| Db | 949 | TGGAGCAATCACTATTTGGAGTATTAATGACCTTGGCTTAATTTGAAGAGTCCAAAGTTT | 1008 | | |
| QY | 361 | CTAAGTCTTACCTTTGAGAGACGCCCTGCAGAAATGTTTTCCGTGTTTAAGAAGATGTCGTC | 420 | | |
| Db | 1009 | CTGCTCTCTACTCTTGGAGAGCGCCCTGCAGAAATGTTTTCCGTACTTAAGAAGATGTTCCGC | 1068 | | |
| QY | 421 | CAATCTTTTGGCAANTTGACCAAAAAGCTACATTCGAAGAGCCATAGATGGATTCAT | 480 | | |

[illegible]

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139455.
PR 16-JUN-1999; 99US-0139455.
PR 17-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0138452.
PR 18-JUN-1999; 99US-0138452.
PR 18-JUN-1999; 99US-0138452.
PR 18-JUN-1999; 99US-0138460.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141843.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0143055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0143803.
PR 09-JUL-1999; 99US-0143920.
PR 12-JUL-1999; 99US-0143927.
PR 13-JUL-1999; 99US-0143544.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0144009.
PR 16-JUL-1999; 99US-0144089.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144329.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144333.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144633.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145211.
PR 26-JUL-1999; 99US-0145222.
PR 26-JUL-1999; 99US-0145227.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145389.
PR 02-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147410.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149375.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149375.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156458.
PR 03-OCT-1999; 99US-0157177.
PR 03-OCT-1999; 99US-0157177.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62 0%; Score 346; DB 21; Length 1335;
Best Local Similarity 79 6%; Pred No 1, 3e-97;
Matches 409; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 45 TTGTGACGACGTGCGCAATTGCAATACGTCCTCTATTGACTCGAATTATGCC 104

| Accession | Gene | Species | Protein | Length | Weight | PI | Instability | Alc | Alc2 | Alc3 | Alc4 | Alc5 | Alc6 | Alc7 | Alc8 | Alc9 | Alc10 | Alc11 | Alc12 | Alc13 | Alc14 | Alc15 | Alc16 | Alc17 | Alc18 | Alc19 | Alc20 | Alc21 | Alc22 | Alc23 | Alc24 | Alc25 | Alc26 | Alc27 | Alc28 | Alc29 | Alc30 | Alc31 | Alc32 | Alc33 | Alc34 | Alc35 | Alc36 | Alc37 | Alc38 | Alc39 | Alc40 | Alc41 | Alc42 | Alc43 | Alc44 | Alc45 | Alc46 | Alc47 | Alc48 | Alc49 | Alc50 | Alc51 | Alc52 | Alc53 | Alc54 | Alc55 | Alc56 | Alc57 | Alc58 | Alc59 | Alc60 | Alc61 | Alc62 | Alc63 | Alc64 | Alc65 | Alc66 | Alc67 | Alc68 | Alc69 | Alc70 | Alc71 | Alc72 | Alc73 | Alc74 | Alc75 | Alc76 | Alc77 | Alc78 | Alc79 | Alc80 | Alc81 | Alc82 | Alc83 | Alc84 | Alc85 | Alc86 | Alc87 | Alc88 | Alc89 | Alc90 | Alc91 | Alc92 | Alc93 | Alc94 | Alc95 | Alc96 | Alc97 | Alc98 | Alc99 | Alc100 | Alc101 | Alc102 | Alc103 | Alc104 | Alc105 | Alc106 | Alc107 | Alc108 | Alc109 | Alc110 | Alc111 | Alc112 | Alc113 | Alc114 | Alc115 | Alc116 | Alc117 | Alc118 | Alc119 | Alc120 | Alc121 | Alc122 | Alc123 | Alc124 | Alc125 | Alc126 | Alc127 | Alc128 | Alc129 | Alc130 | Alc131 | Alc132 | Alc133 | Alc134 | Alc135 | Alc136 | Alc137 | Alc138 | Alc139 | Alc140 | Alc141 | Alc142 | Alc143 | Alc144 | Alc145 | Alc146 | Alc147 | Alc148 | Alc149 | Alc150 | Alc151 | Alc152 | Alc153 | Alc154 | Alc155 | Alc156 | Alc157 | Alc158 | Alc159 | Alc160 | Alc161 | Alc162 | Alc163 | Alc164 | Alc165 | Alc166 | Alc167 | Alc168 | Alc169 | Alc170 | Alc171 | Alc172 | Alc173 | Alc174 | Alc175 | Alc176 | Alc177 | Alc178 | Alc179 | Alc180 | Alc181 | Alc182 | Alc183 | Alc184 | Alc185 | Alc186 | Alc187 | Alc188 | Alc189 | Alc190 | Alc191 | Alc192 | Alc193 | Alc194 | Alc195 | Alc196 | Alc197 | Alc198 | Alc199 | Alc200 | Alc201 | Alc202 | Alc203 | Alc204 | Alc205 | Alc206 | Alc207 | Alc208 | Alc209 | Alc210 | Alc211 | Alc212 | Alc213 | Alc214 | Alc215 | Alc216 | Alc217 | Alc218 | Alc219 | Alc220 | Alc221 | Alc222 | Alc223 | Alc224 | Alc225 | Alc226 | Alc227 | Alc228 | Alc229 | Alc230 | Alc231 | Alc232 | Alc233 | Alc234 | Alc235 | Alc236 | Alc237 | Alc238 | Alc239 | Alc240 | Alc241 | Alc242 | Alc243 | Alc244 | Alc245 | Alc246 | Alc247 | Alc248 | Alc249 | Alc250 | Alc251 | Alc252 | Alc253 | Alc254 | Alc255 | Alc256 | Alc257 | Alc258 | Alc259 | Alc260 | Alc261 | Alc262 | Alc263 | Alc264 | Alc265 | Alc266 | Alc267 | Alc268 | Alc269 | Alc270 | Alc271 | Alc272 | Alc273 | Alc274 | Alc275 | Alc276 | Alc277 | Alc278 | Alc279 | Alc280 | Alc281 | Alc282 | Alc283 | Alc284 | Alc285 | Alc286 | Alc287 | Alc288 | Alc289 | Alc290 | Alc291 | Alc292 | Alc293 | Alc294 | Alc295 | Alc296 | Alc297 | Alc298 | Alc299 | Alc300 | Alc301 | Alc302 | Alc303 | Alc304 | Alc305 | Alc306 | Alc307 | Alc308 | Alc309 | Alc310 | Alc311 | Alc312 | Alc313 | Alc314 | Alc315 | Alc316 | Alc317 | Alc318 | Alc319 | Alc320 | Alc321 | Alc322 | Alc323 | Alc324 | Alc325 | Alc326 | Alc327 | Alc328 | Alc329 | Alc330 | Alc331 | Alc332 | Alc333 | Alc334 | Alc335 | Alc336 | Alc337 | Alc338 | Alc339 | Alc340 | Alc341 | Alc342 | Alc343 | Alc344 | Alc345 | Alc346 | Alc347 | Alc348 | Alc349 | Alc350 | Alc351 | Alc352 | Alc353 | Alc354 | Alc355 | Alc356 | Alc357 | Alc358 | Alc359 | Alc360 | Alc361 | Alc362 | Alc363 | Alc364 | Alc365 | Alc366 | Alc367 | Alc368 | Alc369 | Alc370 | Alc371 | Alc372 | Alc373 | Alc374 | Alc375 | Alc376 |
|-----------|------|---------|---------|--------|--------|----|-------------|-----|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|-----------|------|---------|---------|--------|--------|----|-------------|-----|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|

PA (DUP0) DU PONT DE NEMOURS & CO E. I.
PI Falco SC, Fomodu LO;
XX WPI: 2000-182429/16.
DR P-PSDB: MAF44741.
XX
PM Novel tetrahydrofolate metabolism enzyme used to alter the level of
PM tetrahydrofolate metabolism in plants and seeds
PS
XX Claim 3, Pages 34-35; 37pp; English.
XX
CC The present sequence is a cDNA clone wlm96.pk047.14 encoding
CC 5,10-methylenetetrahydrofolate reductase (MHRF). The clone was
CC isolated from a Wlm96 cDNA library which was prepared using wheat
CC seedlings obtained 96 hrs. after inoculation with E. graminis.
CC MHRF plays a role in the synthesis of methionine.
CC The present sequence is used in the
CC construction of a chimeric gene to alter the level of tetrahydrofolate
CC metabolism enzymes in plants. The enzyme may provide target to facilitate
CC design and/or identification of inhibitors that may be useful as
CC herbicides. The polynucleotide is also useful as a source of probes for
CC genetically and physically mapping the genes and as markers for traits
CC linked to the genes.
XX
S0 Sequence 451 BP; 135 A; 89 C; 101 G; 117 T; 9 other;
Query Match 25.0%; Score 139.6; DB 21; Length 451;
Best Local Similarity 74.0%; Pred. No. 1.8e-31;
Matches 213; Conservative 0; Mismatches 70; Indels 5; Gaps 3;
OY 1 TGATTTGTCACCCATTTTATTAGATACGGATATATTCCTCAAAATTTGGAAGACATGTC 60
DB 163 TTATTAATCCACCCATTTTATTAGATACGGATATATTCCTCAAAATTTGGAAGACATGTC 222
OY 61 GCCAAATTTGGAATTAACGTCGTCTATTGTATGTCGGAATTAATGCCCATTAAATTAACAGA 120
DB 223 GTGAGATTGTGATTAACCTGCCCTATCGTTCTGTGCATTAATGCAATTAATTAACACAAAG 282
OY 121 GCTTTATCCGATGACATGAGGTTTTCGAAACAAAGATACCGATGACATTAATGAGCTG 179
DB 283 GATTTTGTGCGATGACATGAGTTCTGTGCAAAACATAAATTCACCTGAGATTGCTGCTGCT 342
OY 180 TTGAGACCTATCAGCAGCAATGGAAGAGCTGCA--GGCTTAATGGAATTCACCTGGGAGAC 238
DB 343 TGGGTCCTCACTAAAGACATGAGGAGGTTGTGAAAGACATTAATGAGATCCACCTGGTAC 402
OY 239 TGAAT--ATGTCGCAAAAGATTCTTACCTCATGAGCAATTAACATCTGCA 283
DB 403 TGAAGATGTTCAAAATAATTTTNGCTAGTGGGATTAATTAACATCTGCA 450
RESULT 7
AAC91206
ID AAC91206 standard; DNA: 2187 BP.
AAC91206;
XX 20-MAR-2001 (first entry)
DE Human methylenetetrahydrofolate reductase gene SEQ ID NO: 1.
XX
XX Human; schizophrenia; developmental disorder; spina bifida cystica;
XX Tourette's syndrome; bipolar illness; autism; conduct disorder;
XX attention deficit disorder; obsessive compulsive disorder;
XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.
OS Homo sapiens.
PN W0200071754-A1.
XX
XX 30-NOV-2000.
PD


```
OY 298 ATATGAGAAATCTGCATTTGGCAATTAATGAACTTGGGCTTAATTAAGAGATCCAAAG 357
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 963 ACCGCGAGATGGCTACACACAGAGTGTGAAGCGCTGGGATGTGACACAGAGAGCCCA 1042
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 358 TTTTACAGTCTTACCTAGCTGAGAGCGCCCTGCAAAATCTTTCCGCTTAAGAAGATGCC 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1043 GGGCTC---CCCTACCTGGGCTCTCAGTCCGCCACAGCGCCGAGAGGAGATGTAC 1099
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 418 GTCCATCTTTTGGGCAATGACCAAAAGCTACATATCAAGACCATGAGATGGGATC 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1100 GTCCATCTTTGAGGCTCCACAGACCAAGATTAATCTTACCGTACCCAGAGAGATGTAC 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 478 AATACCCACATGGGCGTTGGGCTGATTCGTATTCATCATATGTCATATTCGANT 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1160 AGTTCCTTACAGCGCCGCTGGGCAATTCCTCTCCCTGCTTGGGAGCTGAAGGACT 1219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 538 ATCAGTTCAT 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1220 ACTACCTCTT 1229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AAD20463
ID AAD20463 standard; cDNA; 2219 BP.
XX
XX AAD20463:
XX
XX 03-JAN-2002 (first entry)
XX
DE Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #2.
XX
XX Human methylenetetrahydrofolate reductase (MTHFR) gene therapy; glioma;
KM antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
KM pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;
KM neuroblastoma; leukaemia; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 13..1983
FT /*tag= a
FT /product= "Human methylenetetrahydrofolate reductase
FT (MTHFR) protein"
FT /EC_number= "1.5.1.20"
XX
XX US2001025030-A1.
XX
XX 27-SEP-2001.
XX
XX 01-DEC-2000; 2000US-0728910.
XX
XX 01-MAR-1999; 99US-0258928.
XX
XX (ROZEN/) ROZEN R.
XX (SEKH/) SEKHON J.
XX
XX Rozen R, Sekhon J;
XX
XX WPI: 2001-638509/73.
XX P-PSDB; AAE12607.
XX
XX New antisense nucleic acids, which are methylenetetrahydrofolate
PT reductase inhibitors, useful for treating, stabilizing or preventing
PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
PT neuroblastoma
XX
XX Disclosure: Fig 6; 68bp; English.
XX
XX The invention relates to a non allele-specific antisense nucleic acids,
CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
CC 5-methylenetetrahydrofolate, a co-substrate for methylation of
```

```
CC homocysteine to methionine. The invention provides potential therapy for
CC individuals with MTHFR deficiency. The non allele-specific antisense
CC nucleic acids are useful for treating, stabilizing or preventing cancer,
CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,
CC neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
CC therapy. The present sequence is human methylenetetrahydrofolate
CC reductase (MTHFR) protein cDNA #2. Human MTHFR gene is mapped to
CC chromosome 1p36.3.
XX
XX Sequence 2219 BP: 501 A; 657 C; 620 G; 441 T; 0 other:
XX
XX Query Match 20.9%; Score 116.4; DB 22; Length 2219;
XX Best Local Similarity 53.3%; Pred. No. 2,6e-24;
XX Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
XX
OY 1 TGAATGTACCCCAATTAATTTATGATAGGATATATTCCTCAAAATTTGTGAGCAGCTGC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 683 TCATCATACCGACGACCTTTCTTTGAGGCTGACACATCTTCCGCTTGTGAGGATGCA 742
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 GCCAAATTTGAAATACGCTGTCTATTTGACCTGGAAATATGCCCCATTAATTAACAAG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 743 CCGACATGGGCGATCATCTTGGCCCATGTGCCGGGATCTTCCATCCAGGAGCTACCACT 802
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 121 GCTTATCCGCGATGCTGGTGTGGCAAAACAAGATPACAGCTGACATATGCTGCTT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 803 CCCTTGGGAGCTTGTGAGCTGTGCCAAGCTGAGGTGCCAGAGATCAAGAGACGTGA 862
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 181 TGNAGCTATCAGAGCAATGAGAAAGGCTTATGGAATTCACCGGGAACTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 863 TTGAGCCATCAAGAGCAAGATGCTGCCATCCGCACTATGSCATCAAGGCGCTGGA 922
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 241 AATATGCAAAAGATTTTACCTCATGAGATTAAGACAT---TGCACTTTATPACTTA 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 923 GCTGTGCGAGAGCTTCTGCGAGGTGTGTGTCAGAGGCTTCATCTTACACCTCA 982
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 298 ATATGAGAAATCTGCATTTGGCAATTAATGAACTTGGGCTTAATTAAGAGATCCAAAG 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 983 ACCGCGAGATGGCTACACACAGAGTGTGAAGCGCTGGGATGTGACACAGAGAGCCCA 1042
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 358 TTTTACAGTCTTACCTAGCTGAGAGCGCCCTGCAAAATCTTTCCGCTTAAGAAGATGCC 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1043 GGGCTC---CCCTACCTGGGCTCTCAGTCCGCCACAGCGCCGAGAGAGATGTAC 1099
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 418 GTCCATCTTTTGGGCAATGACCAAAAGCTACATATGAGACCATGAGATGGGATC 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1100 GTCCATCTTTGAGGCTCCACAGACCAAGATTAATCTTACCGTACCCAGAGAGATGTAC 1159
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 478 AATACCCACATGGGCGTTGGGCTGATTCGTATTCATCATATGTCATATTCGANT 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1160 AGTTCCTTACAGCGCCGCTGGGCAATTCCTCTCCCTGCTTGGGAGCTGAAGGACT 1219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 538 ATCAGTTCAT 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1220 ACTACCTCTT 1229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AAT09694
ID AAT09694 standard; cDNA; 2220 BP.
XX
XX AAT09694:
XX
XX 15-OCT-1996 (first entry)
XX
XX Human methylene-tetrahydrofolate-reductase cDNA.
XX
XX Methylenetetrahydrofolate-reductase; MTHFR; gene therapy;
KM cardiovascular disease; neurological disease; folic acid metabolism;
KM EC-1.5.1.20; enzyme; ss.
XX
XX Homo sapiens.
XX
```

| FT | Key | Location/Qualifiers |
|-----------------------|---|---------------------|
| Ft | mat_peptide | 1..1980 |
| Ft | | /tag= a |
| Pn | MO9533054-A1. | |
| Pd | 07-DEC-1995. | |
| Pp | 25-MAY-1995; | 95MO-CA00314. |
| Px | 26-MAY-1994; | 94GB-0010620. |
| Pr | (UIMC-) UNIV MCGILL. | |
| Pa | Goyette P., Rozen R; | |
| Pi | WPI; 1996-030565/03. | |
| Dx | P-PSDB; AAT09694. | |
| Xx | | |
| Pt | Human methylene:tetra:hydro:folate reductase cDNA probe - for detection of sequence abnormalities in methylene:tetra:hydro:folate reductase e.g. in cardiovascular, neurological or folic acid metabolism disorders | |
| Pt | Disclosure: Fig.1A-1F; 66pp; English. | |
| Cc | This sequence encoding human MTHFR has been localized to chromosome 1p36.3. It may be used for the construction of DNA probes which may be used for the identification of sequence abnormalities in patients with severe or mild MTHFR deficiency. The resulting probe may also be used in gene therapy to produce the MTHFR protein. | |
| Sq | Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other; | |
| Query Match | 20.9%; Score 116.4; DB 17; Length 2220; | |
| Best Local Similarity | 53.3%; Pred. No. 2.0e-24; | |
| Matches 293; | Conservative 0; Mismatches 251; Indels 6; Gaps | |
| Oy | 1 TGAATTGCACCCAAATTAATTGATGATGAGATATATCTGTCAAATTTGGTAGACTGTC 60 | |
| Dd | 683 TATCATCATGACGAGCTTCTTGAGAGGTGACACATCTTCTGTGTAGAGCAATGCA 742 | |
| Oy | 61 GCCAAATTTGGATATACGTGTCCTATTGTATGACCTGGAATTTAGGCCATTAAATTACAG 120 | |
| Dd | 743 CGCAATGGGACATCACTTGCCCCATGTCGCCGGGATCTTTCCATCCAGGCTTACCCT 802 | |
| Oy | 121 GCCTTATCCGATGACTGGGTTTTTGCAAAAAGAAGATACACAGCTACATTTATGGCTGTT 180 | |
| Dd | 803 CCCTTTCGGCAGCTTGTAAGGTGTCCAMACGTGAGGTGGCACAGAGATACAAGAGATGA 862 | |
| Oy | 181 TAGAGCCTATCAAGAGCATAGTAGAGAGCTGCATAGAGCTATGATTTACCTGGAGATG 240 | |
| Dd | 863 TTGAGCCCAATCAAAACAGACATGCTGCCATTCGCCAATGTGATGATGACCTGCCGTGA 922 | |
| Oy | 241 AAATGTCCAAAAAGATTTTACATCAATGAAATTAAGACAT--TGACCTTTATATACCTPA 297 | |
| Dd | 923 GGCTGTGCGAGGAGCTTGTGCCAGAGTGCTGTGGTGTGACAGGCTCCACATCTCTACACCTCA 982 | |
| Oy | 298 AATATGAGAAATCTGCATTTGCAATCTATTAAGAACCTGGCCCTATTTAGAAATCCAAG 357 | |
| Dd | 983 ACCGGAGATGGGTACACAGAGAGTCTTAAGCCCTGGGATGTGGAGCTGAG--GATC 1039 | |
| Oy | 358 TTTTCAGTCTTACCTTGTGAGAGAGCCGCTCAAGATTTTCCGTTTAAGAAAGATGTC 417 | |
| Dd | 1040 CCAAGCGTCCCTTACCTCGGGCTCTAGTGTGCCCAACCAAGCCGAGAGAGAGATGTAC 1099 | |
| Oy | 418 GTCCCAATCTTTTGGCAATTCGACCAAAAAGCTCATATCAAGACCATATGATGGGATC 477 | |
| Dd | 1100 GTCCCAATCTTTTGGGCTCCAGACCAAAAGATTCATATCCCTACCCAGAGATGGAGAC 1159 | |
| Oy | 478 AATACCCACATGGAGGCTTGGGGTGAATCTGTGAATCCATATATAGTGTCAATTATCTGATT 537 | |
| Dd | 1160 AGTTCCTTAACAGGCGCTGGGGAGAAATTCCTCTTCCCCTGCTTTGGGAGCTCAAGACT 1219 | |

| QY | 538 | ARCAGTTCAT | 547 |
|-----------|--|--------------------------|------|
| | 1 | 1 | 1 |
| Db | 1220 | ACTACCTCTT | 1229 |
| RESULT 11 | | | |
| ID | AAAS0633 | standard; cDNA: 2220 BP. | |
| XX | AAAS0633; | | |
| XX | 19-DEC-2000 | (first entry) | |
| XX | Human methylenetetrahydrofolate reductase cDNA. | | |
| XX | | | |
| XX | Methylmenetetrahydrofolate reductase; MTHFR; human; folic acid; | | |
| XX | cardiovascular disorder; cancer; neuroblastoma; | | |
| XX | colorectal carcinoma; osteoporosis; neural tube | | |
| XX | neurological disorder; gene therapy; diagnosis; chromosome 1p36.3; | | |
| XX | ss. | | |
| XX | | | |
| XX | Homo sapiens. | | |
| XX | | | |
| XX | Key | Location/Qualifiers | |
| XX | CDS | 13..1983 | |
| XX | | /tag= a | |
| XX | | /EC_number= "1.5.1.20" | |
| XX | mutation | /replace(167,A) | |
| XX | | /tag= b | |
| XX | mutation | /replace(482,A) | |
| XX | | /tag= c | |
| XX | mutation | /replace(559,T) | |
| XX | | /tag= d | |
| XX | mutation | /replace(677,T) | |
| XX | | /tag= e | |
| XX | mutation | /replace(692,T) | |
| XX | | /tag= f | |
| XX | mutation | /replace(764,T) | |
| XX | | /tag= g | |
| XX | mutation | /replace(792+1,A) | |
| XX | | /tag= h | |
| XX | mutation | /replace(985,T) | |
| XX | | /tag= i | |
| XX | mutation | /replace(1015,T) | |
| XX | | /tag= j | |
| XX | mutation | /replace(1018,T) | |
| XX | | /tag= k | |
| XX | mutation | /replace(1298,C) | |
| XX | | /tag= l | |
| XX | mutation | /replace(1317,C) | |
| XX | | /tag= m | |
| XX | | | |
| XX | WO200052205-A2. | | |
| XX | | | |
| XX | 08-SEP-2000. | | |
| XX | | | |
| XX | 28-FEB-2000; 2000WO-1B00442. | | |
| XX | | | |
| XX | 01-MAR-1999; 99US-0258928. | | |
| XX | | | |
| XX | (UIMC-) UNIV MCGILL. | | |
| XX | | | |
| XX | Rozen R, Goyette P; | | |
| XX | | | |
| XX | WPI; 2000-572192/53. | | |
| XX | | | |
| XX | P-PSDB; AAY96186. | | |
| XX | | | |
| XX | CDNA probe for the human methylmenetetrahydrofolate reductase (MTHFR), | | |
| XX | useful in gene therapy and for diagnosing or treating MTHFR deficiency | | |
| XX | which is associated with cardiovascular disorders or cancer | | |
| XX | Claim 1; Fig 1A-F; 93pp; English. | | |

XX The present sequence is that of cDNA coding for human
CC methylenetetrahydrofolate reductase (MTHFR; see AA96186), an enzyme
CC catalysing the NADPH-linked reduction of 5,10-methylenetetrahydrofolate
CC to 5-methyltetrahydrofolate, a co-substrate for methylation of
CC homocysteine to methionine. To identify human MTHFR cDNA, PCR
CC primers based on the porcine sequence were used to screen a human
CC liver lambda-gli0 cDNA library by PCR. A 1266 bp fragment was
CC obtained, and this was used to screen a human colon carcinoma cDNA
CC library to obtain the 2.2 kb clone. The MTHFR gene (see AA50534)
CC maps to chromosome 1p36.6. A cDNA probe for human MTHFR, which
CC hybridises to the present sequence, is claimed. This probe can be
CC used to identify MTHFR sequence abnormalities in individuals with
CC severe or mild MTHFR deficiency. These abnormalities may comprise
CC a mutation selected from 167G to A, 482G to A, 559C to T, 677C to
CC T, 692C to T, 764C to T, 792+1G to A, 985C to T, 1015C to T, 1081C
CC to T, 1298A to C and 1317T to C. MTHFR deficiency may be associated
CC with a cardiovascular disorder, cancer (especially neuroblastoma or
CC colorectal carcinoma), osteoporosis, neural tube defect in an
CC offspring of a patient, neurological disorders, and other disorders
CC influenced by folic acid metabolism. Also claimed are methods for
CC treating MTHFR deficiency by gene therapy or by administration of
CC MTHFR protein. Cancer can be treated by inhibiting MTHFR gene
CC expression or MTHFR protein activity, or by administering an agent
CC that modifies MTHFR gene expression.
XX

Sequence 2220 BP: 501 A; 658 C; 620 G; 441 T; 0 other;

Query Match 20.9%; Score 116.4; DB 21; Length 2220;
Best Local Similarity 53.3%; Pred. No. 2.6e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

QY 1 TGATTGTACCCCAATTTATTTATGATACGATATATCTTCATTAATTTTGAAACGACTGTC 60
DB 683 TCATCATACGACGCTTTCTTTGAGCGTGACACATTCCTTCCCTTGAAGAGATGATCA 742
QY 61 GCCAAATGTGAATACGCTGTCCTATTTGCTACCTGGAATTTGCCCATTAATTAATCAAG 120
DB 743 CCACATGTGGGACATTCACCTTGGCCCATCTGCTCCCGGATCTTTCCATCCAGGCTACCCACT 802
QY 121 GCTTTATCCGATGACTGGTGTTCGCAAAACAAAGATACGCTGACATTTATGGTGGCTT 180
DB 803 CCCTTGGGACGCTTGTGAAGGTGTCCAGCTGAGAGTGTCCACAGAGATCAAGACGTA 862
QY 181 TAGAGCTATTCAGAGACATTAAGAGCTGTCAAGCTTAAAGATTCACCTGGAGACTG 240
DB 863 TTGAGCCATCAAAACAGACATGCTGCCATCCGAACTATGAGATGAGCTGGCGGTGA 922
QY 241 AATATGCAAAAGATTTTACCTCATGGAATTAAGACAT--TGCATCTTTATTAACATTA 297
DB 923 GCGTGTGCGAGAGCTTGTGCGCAGTGGCTGTGTGTCGACGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATTCGATGCGAATTCATTAAGACCTTGGCTATTTGAAGAGTCCAAAG 357
DB 983 ACCGCGAATGGCTACACAGAGTGTGACAGCGCTGGGATGTGACACTGAGAGCCCA 1042
QY 358 TTTCTAGGCTTACCTTGTGAAGAGCGCTGCAATGTTTTCGTTAANGAGATGTC 417
DB 1043 GGGGTC--CCCTACCTGGGCTGTCAAGTCCACCCAGCGCGAGAGAGATGTAC 1099
QY 418 GTCCAACTTTTGGCAAAATGACCAAAAAGCTCAATATCAAGACCATATGATGGATTC 477
DB 1100 GTCCCACTTTTGGGCTTCCAGAGCAAAAGATTAATCATCTACCGTACCAAGAGGGAGC 1159
QY 478 AATACCCACATGGCGTGTGGGATGATTCGTATGATTCATATGATGATTCGATTTATC 537
DB 1160 AGTTCCCTTAAGCGCGCTGGGCAATCTCTCCCTTCCCTTGGGAGACTGAAGACT 1219
QY 538 ATCAGTTCAT 547
DB 1220 ACTACGCTTT 1229

RESULT 12
ID AAD20462 standard; cDNA: 2220 BP.
XX
AC AAD20462;
XX
DT 03-JAN-2002 (first entry)
XX
DE Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #1.
XX
KW Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
KW antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
KW pancreas; lung; brain; cytostatic; colon carcinoma; breast; colorectal;
KW neuroblastoma; leukaemia; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..1983
FT CDS
FT /*tag= a
FT /product= "Human methylenetetrahydrofolate reductase
FT /note= "CDS does not include start codon"
FT /EC.number= "1.5.1.20"
FT /partial
XX
PN US2001025030-A1.
XX
PD 27-SEP-2001.
XX
PF 01-DEC-2000; 200005-0728910.
XX
PR 01-MAR-1999; 9905-0258928.
XX
PA (ROZE/) ROZEN R.
XX (SEKH/) SEKHON J.
XX
PI Rozen R, Sekhon J;
XX
DR WPI: 2001-638509/73.
XX P-PSDB: AAE12606.
XX
PT New antisense nucleic acids, which are methylenetetrahydrofolate
PT reductase inhibitors, useful for treating, stabilizing or preventing
PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
PT neuroblastoma -
XX
PS Disclosure: Fig 1; 68pp; English.
XX
XX The invention relates to a non allele-specific antisense nucleic acids,
XX which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
XX nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
XX catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
XX 5-methylenetetrahydrofolate, a co-substrate for methylation of
XX homocysteine to methionine. The invention provides potential therapy for
XX individuals with MTHFR deficiency. The non allele-specific antisense
XX nucleic acids are useful for treating, stabilizing or preventing cancer,
XX particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
XX lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,
XX neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
XX therapy. The present sequence is human methylenetetrahydrofolate
XX reductase (MTHFR) protein cDNA #1. Human MTHFR gene is mapped to
XX chromosome 1p36.3.
XX
SQ Sequence 2220 BP: 501 A; 658 C; 620 G; 441 T; 0 other;
Query Match 20.9%; Score 116.4; DB 22; Length 2220;
Best Local Similarity 53.3%; Pred. No. 2.6e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

QY 1 TGATTGTACCCCAATTTATTTATGATACGATATATCTTCATTAATTTTGAAACGACTGTC 60
DB 683 TCATCATACGACGCTTTCTTTGAGCGTGACACATTCCTTCCCTTGAAGAGATGATCA 742

| | | | |
|----|------|--|------|
| OY | 61 | GGAATTTGGAAATACGTCCTCTTGTACCGGAATTATGCCAATTAAATTAACAAG | 120 |
| Dd | 743 | CCGACATBGGCATCACTTGCCCCATGGTCCCGGGATCTTTCATTCAGGGCTACACT | 802 |
| OY | 121 | GCTTATCCGCGTAGCTGGTGTGTTGGCAAAAACAATACGCAGCTCATTATAGCTGCT | 180 |
| Dd | 803 | CCCTTCGGCACCTTGGAAAGCTGTCCACAGCTGGAGGTGCCAAGAGATCAAAGACGTGA | 862 |
| OY | 181 | TGAGACCTTCACAGGACANTAGAAGACCTGTCAAGCTTATGGAATTCACCTGGAACTG | 240 |
| Dd | 863 | TTGACCACTAACAAAGACACAGCTGTCTCCATCCGACACTATGGATCGAGCTGGCGCTGA | 922 |
| OY | 241 | AAATGTGCMAAAGATTATTAGCTATGGAATTAGACAT---TGCACTTTTATACATAA | 297 |
| Dd | 923 | GCCTGTACGAGAGCTCTGGCACATGGCTGGTGGACAGGCTCCACTCTTACACCTCA | 982 |
| OY | 298 | AATAGSAGAAATCTGCATTGGCAATTACTAATGAACCTTGGCTAATTGAAGATCCAAAG | 357 |
| Dd | 983 | ACCAGCAATAGGTCTACACAACAAGGTGTGTAAGCGCTGGAGATGTGAGCTAGAGCCCA | 1044 |
| OY | 358 | TTTTAGGTCTTACCTTGTSGAGAGCCCTGCAAAATGTTTTCGTTTAAAGAGATGTCC | 417 |
| Dd | 1043 | GGCGTC---CCCTTACCCTGGGGCTCTCGCTCCCACCACCAACGCGCAGAGGAGATGTAC | 1099 |
| OY | 418 | GTCACATCTTTGGGCAATGACCAAAAAAGCTACATATCAAGAGACATAGATGGATC | 477 |
| Dd | 1100 | GTCCCATCTTCTGGGCTCTCAGACCAAAAGATTTACATCTTACCGTACCCAGGAGTGGAGC | 1155 |
| OY | 478 | AATACCACATATGGCCCTTGGGGTATTTCTGTAAATCCATCATATGTGTGATTAATCATTT | 537 |
| Dd | 1160 | ACCTTCCCTCAAGCGCCGCTGGGGCAATTCCTCTCTCCCTTGCGCTTGGGAGCTGAAGTACT | 1211 |
| OY | 538 | ATCAGTTCAT | 547 |
| Dd | 1220 | ACTACCTCTT | 1229 |

| RESULT 13 | ABK13501 | ID | ABK13501 standard; cDNA; 2220 BP. |
|-----------|----------|----|--|
| XX | XX | AC | ABK13501; |
| XX | XX | DT | 09-APR-2002 (first entry) |
| XX | XX | DE | DNA encoding human methyltetrahydrofolate reductase (MTHFR). |
| XX | XX | XX | Methyltetrahydrofolate reductase; MTHFR; neuroleptic; neuroprotective; triamgilliser; nootropic; antidepressant; anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis; anxiety; dementia; depression; epilepsy; Huntington's disease; angriale; demyelinating disease; multiple sclerosis; pain; Parkinson's disease; psychosis; stroke; cardiovascular disorder; cancer; osteoporosis; metabolic disease; endocrine disease; labour error of metabolism; inflammation; immune disorder; human; neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease; chromosome 1p36; ss; Ec number 1.5.1.20. |
| XX | XX | OS | Homo sapiens. |
| XX | XX | EH | Key |
| XX | XX | FT | Location/Qualifiers |
| XX | XX | FT | 1..1983 |
| XX | XX | FT | /*tag a |
| XX | XX | FT | /product= "MTHFR" |
| XX | XX | FT | /note= "Methyltetrahydrofolate reductase" |
| XX | XX | PN | WO200196598-A2. |
| XX | XX | PD | 20-DEC-2001. |
| XX | XX | PD | 12-JUN-2001; 2001WO-CA00867. |

[illegible]

Db 1100 GTCACATCTTGGGCTCCAGACCAAGAGTACATCTACCGTACCCAGAGGAGGACG 1159
QY 478 AATACCCACATGGGGCTGTGATTCCTATCCATCATATGTCATTATCTGATT 537
1160 AGTTCCTTAACGGCGCTGGGCAATTCCTTCCCTCCCTTGGGAGCTGAAGACT 1219
QY 538 ATCAGTTCAT 547
Db 1220 ACTACCTCTT 1229

RESULT 14
ABK13539
ID ABK13539 standard; cDNA; 2220 BP.
XX ABK13539;
XX
DT 09-APR-2002 (first entry)
XX
XX Human methyltetrahydrofolate reductase (MTHFR) G167A allele.
XX
XX Methyltetrahydrofolate reductase; MTHFR; neuroleptic;
XX neuroprotective; tranquilizer; nootropic; antidepressant;
XX anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis;
XX anxiety; dementia; depression; epilepsy; Huntington's disease;
XX migraine; demyelinating disease; multiple sclerosis; pain;
XX Parkinson's disease; psychosis; stroke; cardiovascular disorder;
XX cancer; osteoporosis; metabolic disease; endocrine disease;
XX inborn error of metabolism; inflammation; immune disorder; human;
XX neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease;
XX ss; EC number 1.5.1.20; mutant.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX
XX Key Location/Qualifiers
FT 1..1983
FT /tag= a
FT /product= "MTHFR"
FT /note= "Methyltetrahydrofolate reductase"
FT replace(167,G)
FT /tag= b
XX
XX W0200196598-A2.
XX
XX PD 20-DEC-2001.
XX
XX 12-JUN-2001; 2001WO-CA00867.
XX
XX PF 12-JUN-2000; 2000US-0592595.
XX
XX PR 12-JUN-2000; 2000US-0592595.
XX
XX PA (UYMC-) UNIV MCGILL.
XX
XX PI Rozen R;
XX
XX WPI; 2002-130741/17.
XX
XX DR P-PSDB; AAU5421.
XX
XX
XX Diagnosing subjects at risk for or suffering from a psychosis,
XX particularly schizophrenia comprises determining the presence of a
XX heterozygous methyltetrahydrofolate reductase mutant allele in the
XX subject
XX
XX Claim 4; Page -; 102pp; English.
XX
XX The invention describes a method of diagnosing a psychosis in a subject,
XX or a risk for or propensity to psychosis in a subject comprising
XX determining the presence of a heterozygous methyltetrahydrofolate
XX reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a
XX subject. The method is useful for diagnosing subjects at risk of, or
XX suffering from a psychosis, particularly schizophrenia but includes
XX manic-depressive disease, organic psychotic disorders, psychosis in
XX alcohol or drug intoxication, postinfection psychosis, postpartum

CC psychosis, senile psychosis, traumatic psychosis and acute idiopathic
CC psychotic illnesses. A new pharmaceutical composition is used for
CC treating amyotrophic lateral sclerosis, anxiety, dementia, depression,
CC epilepsy, Huntington's disease, migraine, demyelinating disease, multiple
CC sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or
CC stroke. Deficiency may be associated with diseases including
CC cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal
CC carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors
CC of metabolism, inflammation, immune disorders, neoplastic disease and
CC renal disease. This sequence encodes a mutant human
CC methyltetrahydrofolate reductase (MTHFR, EC number 1.5.1.20), in which
CC nucleotide 167 has been altered to produce a mutant allele, described in
CC the method of the invention.
CC Note: This sequence does not appear in the specification but has been
CC created from the wild type sequence (ABK13501) using information given
CC in claim 4 of the invention.
XX
SQ Sequence 2220 BP; 502 A; 658 C; 619 G; 441 T; 0 other:
Query Match 20.9%; Score 116.4; DB 24; Length 2220;
Best Local Similarity 53.3%; Pred. No. 2,6e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
QY 1 TGATGTCACCCCAATTTATTTATGATGATATATCCCAAAATTTGGAAGCACTGTC 60
Db 683 TCATCATCAGCGAGCTTTCTTTAGAGCTGACATTTCTCCGCTTTGGAAGGATGCA 742
QY 61 GCCAAATTGGAATTAACGTCGCTCTATTGCTGGAATTTATGCCCATTAATTAACAAG 120
Db 743 CCAGCATGAGGACATCTGCCCCCATGTCGCCGGGATCTTCCATCCAGGGATCCACT 802
QY 121 GCTTTATCCGCAATGATGAGTGGTTTGCAAAACAAATACAGCTGACATTAATGCTGCTT 180
Db 803 CCTTCGGCAGCTGTGGAAGCTGTCCAAAGCTGAGGAGTGCACAGAGATCAAGAGCGTGA 862
QY 181 TAAAGCCTATCAAGACATGAAAGAGCTGCAAGCTTAATGGAATTAATCACTGGGAACTG 240
Db 863 TTGAGCCATCAAGCAAGCAAGATGCTGCATCGCAACTATGACATGAGCTGAGCCGTGA 922
QY 241 AATATGCAAAAAGATTTTAAAGCTATGGAATTAAGACAT--TGACATTTATACACTAA 297
Db 923 GCTGTGACAGAGAGCTTCGGGACATGCTGAGGAGGACAGCTCCACTTCACCTCA 982
QY 298 ATATGAGAAATGATGATGCAATACATTAAGAACCTTGAGCTTAATGAAGTGCACAA 357
Db 983 ACCGCGAATGCTACACACAGAGTGTGAAGCCCTGGGAGATGTGACTGAGACCCCA 1042
QY 358 TTCTTAGGTCCTTACCTTGAGAGCGCCCTGCAAAATGTTTCGCTTAAAGATGTCC 417
Db 1043 GAGGTC--CCCTACCTTGAGGCTCAATGATGCCACCCCAAGSCCGAGAGGAAATGTATC 1099
QY 418 GTCCATCTTTGGGCAATGACCAAAAAGCTATATCAAGACCATAGATGATGATC 477
Db 1100 GTCCATCTTTGGGCTGCTCCAGACCAAGATTAATCATCTACCGTACCCAGAGTGGAGCG 1159
QY 478 AATACCCACATGGGCGCTGGGCTGATTCCTGTAATCATATATGTCATTTATCTGATT 537
Db 1160 AGTTCCTTAACGGCGCTGGGCAATTCCTTCCCTCCCTTGGGAGCTGAAGACT 1219
QY 538 ATCAGTTCAT 547
Db 1220 ACTACCTCTT 1229

RESULT 15
ABK13540
ID ABK13540 standard; cDNA; 2220 BP.
XX ABK13540;
XX
XX 09-APR-2002 (first entry)
XX
XX Human methyltetrahydrofolate reductase (MTHFR) G482A allele.
DE

us-09-720-451-5.rni



GenCore version 5.1.3
(c) 1993 - 2003 Compu

2 ; Search time 55 seconds

3111.375 Million cell updates/sec

.....tcagtctcatgcygccacgtg 558

dues

imeters: 882724

2003

redicted by chance to have a score of the result being printed on the normal score distribution.

RES

| | Description |
|----------|--------------------|
| 878-23 | Sequence 23, Appl |
| 448-1 | Sequence 1, Appl |
| 000-3 | Sequence 3, Appl |
| 928-3 | Sequence 3, Appl |
| 000-1 | Sequence 1, Appl |
| 928-1 | Sequence 1, Appl |
| 527-140 | Sequence 140, Appl |
| 207A-54 | Sequence 54, Appl |
| 630-3 | Sequence 3, Appl |
| 017B-931 | Sequence 931, Appl |
| 504-41 | Sequence 41, Appl |
| 266-41 | Sequence 41, Appl |
| 760A-1 | Sequence 1, Appl |
| 163-3 | Sequence 3, Appl |
| 891A-13 | Sequence 13, Appl |
| 285-1 | Sequence 1, Appl |
| 764-1 | Sequence 1, Appl |
| 280-5 | Sequence 5, Appl |
| 783-1 | Sequence 1, Appl |
| 233-5 | Sequence 5, Appl |
| 626-5 | Sequence 5, Appl |
| 307A-5 | Sequence 5, Appl |
| 928-5 | Sequence 5, Appl |
| 137-5 | Sequence 5, Appl |
| 3624-5 | Sequence 5, Appl |
| 017B-846 | Sequence 846, Appl |

Sequence 14, April 1983, April 10, April 3, April 1013, April 1, April 1, April 1, April 96, April 11, April 11, April 11, April 11, April 800, April 800, April 780, April 1011, April 1011, April 14, April

ANALYTES

ch 1971;

6; Gaps 2,

[illegible]

```
QY 358 TTCTAGGTCCTTACCTTGAGAGCCCTGCAAAATGTTTCGGTGTAAAGAGATGCC 417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1031 GGCGTC---CCCTACCGCTGGGCTCTCAGTCCACCCCAAGCGCCGAGAGAGATGAC 1087

QY 418 GTCCATCTTTGGGCAAAATGACCAAAAGCTACATATGCAAGCCATAGATGGGATC 477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1088 GTCCATCTTTGGGCTCTCAGTCCACCCCAAGAGCTTACCTACCGTCCAGAGATGGGAGC 1147

QY 478 AATACCAATGCGGCTTGGGCTGATCCCTGATACATCATATGATGCTATTCGATT 537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1148 AGTTCCTTACGCGCGCTGGGCAATTCCTTCCCTGCGCTTGGGAGCTGAAAGACT 1207

QY 538 ATCAGTTCAT 547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1208 ACTACGCTCT 1217

RESULT 2
US-09-318-448-1
; Sequence 1, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenicos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-1

Query Match      20.9%; Score 116.4; DB 4; Length 2187;
Best Local Similarity 53.3%; Pred. No. 2.1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

QY 1 TGATGTACCCCAATTATTTATGATAGGATATATTCCTGCAAAATTTGTGAGACAGCTGTC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 TCATCATCAGCAGCGCTTTCTTGTGAGGCTGACACATCTTCGCTTGTGAAGGCAATGCA 733

QY 61 GCCAAATTTGCAATACGTGCTCTATGTACTGGAATTAATGCCATTATATTTACAAAG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 734 CCGACATGGGCAATCACTGGCCCATCGTCCCGGATCTTCCCATCCAGGCTACCACT 793

QY 121 GCTTATCCGATGACTGGTGTGTTGCAAAACAAAGATACAGCTGACATTAATGGCTGCT 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 794 CCCTTGGCAGCTTGTGAAGCTGTCCAAAGCTGAGAGGTGCCAAGAGATCAAGAGCTGA 853

QY 181 TAGAGCCTTATCAAGCAATGATGAGAGCTGTCAAGCTTATGAAATCACTGCTGAGAACG 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 TTGAGCCATTCAAAGCAAGCATGCTGCTCGCACTATGAGATGAGACTGAGCTGCCGTGA 913

QY 241 AAATGTGCAAAAAGATTTTACCTCATGTGAATTAAGCAT---TGCATCTTATACACTRA 297
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 GCGTGTGCAAGAGCTTCTGCGCAGTGGCTGTGTCAGAGGCTCACTCTTACACCTCTCA 973

QY 298 ATATGAGAAATTCGATTTGCAATTAATGAACCTTGGCCTTAATGAAGAGCTCCAAAG 357
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 974 ACCCGAGATGGCTACACAGAGGTGCTGAAGGCGCTGGGATGTGACAGTGAAGACCCA 1033

QY 358 TTCTAGGTCCTTACCTTGAGAGCCCTGCAAAATGTTTCGGTGTAAAGAGATGCC 417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1034 GGCGTC---CCCTACCGCTGGGCTCTCAGTCCACCCCAAGCGCCGAGAGAGATGAC 1090

QY 418 GTCCATCTTTGGGCAAAATGACCAAAAGCTACATATGCAAGCCATAGATGGGATC 477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1091 GTCCATCTTTGGGCTCTCAGTCCACCCCAAGAGCTTACCTACCGTCCAGAGATGGGAGC 1150
```

```
QY 478 AATACCAATGCGGCTTGGGCTGATCCCTGATACATCATATGATGCTATTCGATT 537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 AGTTCCTTACGCGCGCTGGGCAATTCCTTCCCTGCGCTTGGGAGCTGAAAGACT 1210

QY 538 ATCAGTTCAT 547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1211 ACTACGCTCT 1220

RESULT 3
US-08-738-000-3
; Sequence 3, Application US/08738000
; Patent No. 6074821
; GENERAL INFORMATION:
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; TITLE OF INVENTION: REDUCTASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: KLAUBER & JACKSON
; STREET: Continental Plaza - 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,000
; FILING DATE:
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CA95/00314
; FILING DATE: 25-MAY-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9410620.0
; FILING DATE: 26-MAY-1994
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2219 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..1983
US-08-738-000-3

Query Match      20.9%; Score 116.4; DB 3; Length 2219;
Best Local Similarity 53.3%; Pred. No. 2.1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

QY 1 TGATGTACCCCAATTATTTATGATAGGATATATTCCTGCAAAATTTGTGAGACAGCTGTC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 TCATCATCAGCAGCGCTTTCTTGTGAGGCTGACACATCTTCGCTTGTGAAGCAATGCA 742

QY 61 GCCAAATTTGCAATACGTGCTCTATGTACTGGAATTAATGCCATTATATTAACAAAG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CCGACATGGGATATCATTCGCCCATCGTCCCGGATCTTCCCATCCAGGCTACCACT 802

QY 121 GCTTATCCGATGACTGGTGTGTTGCAAAACAAAGATACAGCTGCAATTAATGGCTCTT 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 CCCTTGGCAGCTTGTGAAGCTGTCTCAAGCTGAGAGGTGCCACAGAGATCAAGAGCTGA 862

QY 181 TAGAGCCTTATCAAGCAATGATGAGAGCTGCTCAAGCTTATGAAATTCACCTGGAACG 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 863 TTGAGCCATTCAAAGCAAGCATGCTGCAATCCGCAACTGATGCAATCGAGCTGGCCTGA 922
```

```
OY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT---TGCATCTTTATACACTAA 297
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 923 GCCTGTGCCAGAGAGCTTGTGCCAGTGGCTGTGCCAGAGGCTCCTACCTTGTACACCTCA 982
OY 298 ATATGAGAAATCTGCATTGGCAATACATATGACCTTGGCTTAATGGAAGTCCAAAG 357
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 963 ACCGCGAGATGGCTACACAGAGAGTGTGAAGGCTGGAGATGTGACTGAGAGCCCA 1042
OY 358 TTTCTAGTCTTACCTTGGAGAGCGCCCTGCAGAAATGTTTCCGTGTTAAAGAATGTCC 417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1043 GCGCTC---CCCTACCCCTGGGCTCTCAGTGCACCCCAAGCGCCGAGAGAGATGTAC 1099
OY 418 GTCCAACTTTTGGGCAAAATCGAACAAAAGCTACATATCAAGAGCAATAGATGGAGTC 477
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1100 GTCCCACTTTTGGGCTCTCCAGACCAAAAGATTACATCTACCGTACCCAGAGATGGAGC 1159
OY 478 AATACCAATGAGGCGTTGGGGTGTATCTCATCATATATGATGTGATTTATCTGATT 537
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1160 AGTTCCTTAAGCGCGCTGGGCAATTCCTCTTCCCTGGCTTGGGAGCTGAAGACT 1219
OY 538 ATCAGTTCAT 547
      | | | | |
DB 1220 ACTACCTCTT 1229
```

RESULT 4

```
US-09-258-928-3
; Sequence 3, Application US/09258928
; Patent No. 6218120
; GENERAL INFORMATION:
; APPLICANT: KOZEN, Rima
; APPLICANT: GOVETTE, Philippe
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; TITLE OF INVENTION: REDUCTASE
; FILE REFERENCE: 04844/005002
; CURRENT APPLICATION NUMBER: US/09/258,928
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1983)
US-09-258-928-3
```

```
Query Match 20.9%; Score 116.4; DB 4; Length 2219;
Best Local Similarity 53.3%; Pred. No. 2.1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
```

```
OY 1 TGATGTACCCCAATTTATATAGCATAGCATATATTCCTCAAAATTTGTGAAGACTGTC 60
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 663 TCATCATCAGCAGCGCTTTCTTGTAGGCTGCACATCTTCCGCTTGTGAAGCATGCA 742
OY 61 GCCAAATTTGAATAGTGTCTTGTACTGGAATATGATGCGCAATTAATTAATTAAGA 120
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 743 CCGACATGGGCACTATGCCCCCATCGTCCCGGGATCTTCCATCCAGGGCTTACACT 802
OY 121 GCTTATCCGATGACTGGGTTTGGCAAAACAAGATACAGAGTCAAGATTAATGAGCTGTT 180
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 803 CCTTGGGAGCTTGTGAAGCTGTCCAAGCTGAGAGTGCCAAAGAGATCAAGAGAGTGA 862
OY 181 TAGAGCTATCAGAGCATGAGAGAGTGTCAAGAGCTTAAGAGCTTACCTGAGAGAGT 240
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 863 TTAGGCAATCAAAAGCAACGATGCTGCATCCGCAACTATAGCATGAGACTGGCCGTGA 922
```

```
OY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT---TGCATCTTTATACACTAA 297
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 923 GCCTGTGCCAGAGAGCTTGTGCCAGTGGCTGTGCCAGAGGCTCCTACCTTGTACACCTCA 982
OY 298 ATATGAGAAATCTGCATTGGCAATACATATGACCTTGGCTTAATGGAAGTCCAAAG 357
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 963 ACCGCGAGATGGCTACACAGAGAGTGTGAAGGCTGGAGATGTGACTGAGAGCCCA 1042
OY 358 TTTCTAGTCTTACCTTGGAGAGCGCCCTGCAGAAATGTTTCCGTGTTAAAGAATGTCC 417
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1043 GCGCTC---CCCTACCCCTGGGCTCTCAGTGCACCCCAAGCGCCGAGAGAGATGTAC 1099
OY 418 GTCCAACTTTTGGGCAAAATCGAACAAAAGCTACATATCAAGAGCAATAGATGGAGTC 477
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1100 GTCCCACTTTTGGGCTCTCCAGACCAAAAGATTACATCTACCGTACCCAGAGATGGAGC 1159
OY 478 AATACCAATGAGGCGTTGGGGTGTATCTCATCATATATGATGTGATTTATCTGATT 537
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1160 AGTTCCTTAAGCGCGCTGGGCAATTCCTCTTCCCTGGCTTGGGAGCTGAAGACT 1219
OY 538 ATCAGTTCAT 547
      | | | | |
DB 1220 ACTACCTCTT 1229
```

RESULT 5

```
US-08-738-000-1
; Sequence 1, Application US/08738000
; Patent No. 6074821
; GENERAL INFORMATION:
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; TITLE OF INVENTION: REDUCTASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLABBER & JACKSON
; STREET: Continental Plaza - 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,000
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/CA95/00314
; FILING DATE: 25-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9410620.0
; FILING DATE: 26-MAY-1994
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1980
US-08-738-000-1
```

```
Query Match 20.9%; Score 116.4; DB 3; Length 2220;
Best Local Similarity 53.3%; Pred. No. 2.1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
```

```

QY 1 TGATGTGACCCCAATTTATATGATGAGATATATCTCAATTTGTGACGACTGTC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 TCATCATACGACGCTTTCTTTGAGGCTGACACATCTTCCGCTTTGTGAAGGATGCA 742
QY 61 GCCAATTTGGAAATAGCTGTCTGTGACGTAATTTGCGCATTAATTAATTCAGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CCGACATGCGCATCTTGTGCGCATGTCCCGGATCTTTCCATCCAGGCTACACT 802
QY 121 GCTTATCCGACAGCTGGTGTGCAAAACAAATAGACGCTGACATATAGGCTGCT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 CCTTCGCGACGCTGTGAAGCTGTCCAGCTGAGAGTCCACAGAGATCAGAGCTGA 862
QY 181 TAGAGCTTATCAGACAAATAGAGCTGTCAAGCTTATGAAATTCACCTGGAACTG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 863 TTGAGCCATCAAGACAGACGATGTGCTGCGCACTATGAGCTGAGAGCTGGTGA 922
QY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT--TGCATCTTTATACCTAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 923 GCGTGTGCGAGAGCTTCTGGCCAGTGCTGTGATGCGCAGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATCTGCATTGCAATPACTAATGAACTTGGCTAATTTGAAGTCCAAAG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 983 ACCGCGAGATGCTGTACACAGAGGTGCTGAGCGCTGAGATGTGACTGAGAGACCCA 1042
QY 358 TTTCTAGGCTCTTACCTTGAGAGAGCGCCGCAAAATGTTTCCGTGTTAAAGAAATGTC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 GCGCTC--CCCTACCTGGGCTCTCAGTGCACCCCAAGCGCCGAGAGAAAGATGTAC 1099
QY 418 GTCCAACTTTTGGGCAATGCAACCAAAAGCTATCAAGAGACCATAGATGAGATG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GTCCATCTTCTGGGCTCTCAGACCAAAAGTTACATCTACCGACCGAGAGTGGAGC 1159
QY 478 AATACCCCAATGAGGCTTGGGCTGATTCCTGTAATCCATCATATGAGTATTCGAT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 AGTTCCTCAACGCGCCGTGGGCAATTCCTCTTCCCTGCTTTGGGAGCTGAGAGACT 1219
QY 538 ATCAGTTCAT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 ACTACCTCTT 1229

```

```

RESULT 6
US-09-258-928-1
; Sequence 1, Application US/09258928
; Patent No. 6218120
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005002
; CURRENT APPLICATION NUMBER: US/09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: GB 9410620.0
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
US-09-258-928-1

Query Match 20.98; Score 116.4; DB 4; Length 2220;
Best Local Similarity 53.38; Pred. No. 2,1e-27;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;
QY 1 TGATGTGACCCCAATTTATATGATGAGATATATCTCAATTTGTGACGACTGTC 60

```

```

Db 683 TCATCATACGACGCTTTCTTTGAGGCTGACACATCTTCCGCTTTGTGAAGGATGCA 742
QY 61 GCCAATTTGGAAATAGCTGTCTGTGACGTAATTTGCGCATTAATTAATTCAGG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CCGACATGCGCATCTTGTGCGCATGTCCCGGATCTTTCCATCCAGGCTACACT 802
QY 121 GCTTATCCGACAGCTGGTGTGCAAAACAAATAGACGCTGACATATAGGCTGCT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 803 CCTTCGCGACGCTGTGAAGCTGTCCAGCTGAGAGTCCACAGAGATCAGAGCTGA 862
QY 181 TAGAGCTTATCAGACAAATAGAGCTGTCAAGCTTATGAAATTCACCTGGAACTG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 863 TTGAGCCATCAAGACAGACGATGTGCTGCGCACTATGAGCTGAGAGCTGGTGA 922
QY 241 AATGTGCAAAAAGATTTTACCTCATGGAATTAAGACAT--TGCATCTTTATACCTAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 923 GCGTGTGCGAGAGCTTCTGGCCAGTGCTGTGATGCGCAGGCTCCACTTCTACACCTCA 982
QY 298 ATATGAGAAATCTGCATTGCAATPACTAATGAACTTGGCTAATTTGAAGTCCAAAG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 983 ACCGCGAGATGCTGTACACAGAGGTGCTGAGCGCTGAGATGTGACTGAGAGACCCA 1042
QY 358 TTTCTAGGCTCTTACCTTGAGAGAGCGCCGCAAAATGTTTCCGTGTTAAAGAAATGTC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 GCGCTC--CCCTACCTGGGCTCTCAGTGCACCCCAAGCGCCGAGAGAAAGATGTAC 1099
QY 418 GTCCAACTTTTGGGCAATGCAACCAAAAGCTATCAAGAGACCATAGATGAGATG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GTCCATCTTCTGGGCTCTCAGACCAAAAGTTACATCTACCGTACCGAGATGGAGAGC 1159
QY 478 AATACCCCAATGAGGCTTGGGCTGATTCCTGTAATCCATCATATGAGTATTCGAT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 AGTTCCTCAACGCGCCGTGGGCAATTCCTCTTCCCTGCTTTGGGAGCTGAGAGACT 1219
QY 538 ATCAGTTCAT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1220 ACTACCTCTT 1229

```

```

RESULT 7
US-08-961-527-140/c
; Sequence 140, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

```

TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 140:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28882 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-140

Query Match
 Best Local Similarity 47.8%; Pred. No. 0.012;
 Matches 150; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

QY 4 TTGACCCCAATTATTTATGATAGCGATATATTCCTCAATTGAGACGCTGCGC 63
 DB 15801 TCGTAACGAGCTCTTCTTGACAGAGCGCTCTATGATTTGAGCAAGTATCT 15742
 QY 64 AAATGGATACGCTGCTTATGACGGAATATGCGCATTTATATACAGGGCT 123
 DB 15741 TGGCTGGATGATGTTCCTCATGACGAGATATTCGCAATTCGATCGAATCAGG 15682
 QY 124 TTATCGS---CATGACCTGGCTTTCGCAAAACAAGATACGCTGCAATTCGCTGCT 180
 DB 15681 CTCGCGACTCTGAGAGCTGTGAGAAATTCATCTCCACGCAATTTAAGCATCT 15622
 QY 181 TAGAGCTATCAAGCAATGAGAAAGCTGTCAAGCTTATGATTCACCTGGAGATG 240
 DB 15621 TAGCAAGTATGACATGACCTTACGCTGACGACGAGCTGCTGATGCAATGG 15562
 QY 241 AAATGCAAAAAGATTTTACCTCAGTATGAGATTAAGACATTCGATCTTTATACATA 300
 DB 15561 ACCAATCGTGACTGTGATGAGATGTTGCCGTGCTCATCTATGATGATA 15502
 QY 301 TGGAGAAATCTGCA 314
 DB 15501 ATGCTGATACGCA 15488

RESULT 8
 US-08-858-207A-54
 Sequence 54, Application US/08858207A
 Patent No. 6348328
 GENERAL INFORMATION:
 APPLICANT: Black, Michael
 APPLICANT: Hodgson, John
 APPLICANT: Knowles, David
 APPLICANT: Nicholas, Richard
 APPLICANT: Stodola, Robert
 TITLE OF INVENTION: No. 6348328el Compounds
 NUMBER OF SEQUENCES: 552
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Smithline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: IBM Compatible
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/858,207A
 FILING DATE: 09-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/017670
 FILING DATE: 14-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimml, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P50475

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1449 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-858-207A-54

Query Match
 Best Local Similarity 47.5%; Pred. No. 0.0093;
 Matches 149; Conservative 0; Mismatches 162; Indels 3; Gaps 1;

QY 4 TTGACCCCAATTATTTATGATAGCGATATATTCCTCAATTGAGACGCTGCGC 63
 DB 871 TCGTAACGAGCTCTTCTTGACAGAGCGCTCTATGATTTGAGCAAGTATTT 930
 QY 64 AAATGGATACGCTGCTTATGACGGAATATGCGCATTTATATACAGGGCT 123
 DB 931 TGGCTGGATGATGTTCCTCATGACGAGATATTCGCAATTCGATCGAATCAGG 990
 QY 124 TTATCGS---CATGACCTGGCTTTCGCAAAACAAGATACGCTGCAATTCGCTGCT 180
 DB 991 CTCGCGACTCTGAGAGCTGTGAGAAATTCATCTCCACGCAATTTAAGCATCT 1050
 QY 181 TAGAGCTATCAAGCAATGAGAAAGCTGTCAAGCTTATGATTCACCTGGAGATG 240
 DB 1051 TAGCAAGTATGACATGACCTTACGCTGACGACGAGCTGCTGATGCAATGG 1110
 QY 241 AAATGCAAAAAGATTTTACCTCAGTATGAGATTAAGACATTCGATCTTTATACATA 300
 DB 1111 ACCAATCGTGACTGTGATGAGATGTTGCCGTGCTCATCTATGATGATA 1170
 QY 301 TGGAGAAATCTGCA 314
 DB 1171 ATGCTGATACGCA 1184

RESULT 9
 US-08-288-630-3/c
 Sequence 3, Application US/08288630
 Patent No. 5837848
 GENERAL INFORMATION:
 APPLICANT: ELY, SUSAN
 APPLICANT: EVANS, IAN J
 APPLICANT: SCHUCH, WOLFGANG W
 TITLE OF INVENTION: ROOT-SPECIFIC PROMOTER
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/288,630
 FILING DATE: 10-AUG-1994
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/017670
 FILING DATE: 14-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimml, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P50475

```

PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GB 9006017.9
? FILING DATE: 16-MAR-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: KOKULIS, PAUL N
? REGISTRATION NUMBER: 16,773
? REFERENCE/DOCKET NUMBER: 215964/SEE3569USCIP
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 861-3000
? TELEFAX: (202) 822-0944
? TELEX: 6714627 CUSH
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4203 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
?
US-08-288-630-3

Query Match
? Best Local Similarity 50.3%; Score 35.4; DB 2; Length 4203;
? Matches 87; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 156 ATACCACTGACATATGCGCTTACAGCCTATCAAGACATGAGACATGCTCAG 215
DB 630 ATGCACAGTGATTATTTTGTCAANTTAAGGCAACTTTGACCAAGTAAAGGA 571
QY 216 GCTTAGCAATTCACCTGGAGACTGAAATGCAAAAGATTTAGCTCATGAAATAG 275
DB 570 AGTTTGTCTTCTGCTTACGACGACGACGACGACGACGACGACGACGACGAC 511
QY 276 ACATGCACTTATACACTAATATGAGGAAATGCTGATGCAATTCAT 328
DB 510 TCACTATTTTCAATAGAGAGGAAATGAGAGGATATCTTTGTAGAAATAT 458

RESULT 10
US-09-221-017B-931
? Sequence 931 Application US/09221017B
? Patent No 6444789
? GENERAL INFORMATION:
? APPLICANT: ROSS, Bruce C.
? TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
? NUMBER OF SEQUENCES: 1120
? CORRESPONDENCE ADDRESS: 6 FORSTER
? ADDRESSEE: MORRISON & FORSTER
? STREET: 755 PAGE MILL ROAD
? CITY: PALO ALTO
? STATE: CA
? COUNTRY: USA
? ZIP: 94304-1018
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: Windows
? SOFTWARE: FASTSEQ for Windows Version 2.0b
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/221.017B
? FILING DATE: 23-DEC-1998
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PP1182
? FILING DATE: 31-DEC-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PP1546
? FILING DATE: 30-JAN-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PP2911
? FILING DATE: 09-APR-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/AU98/01023
? FILING DATE: 10-DEC-1998

```

```

ATTORNEY/AGENT INFORMATION:
? NAME: Monroy, Gladys H
? REGISTRATION NUMBER: 32,430
? REFERENCE/DOCKET NUMBER: 27340-20021.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-813-5600
? TELEFAX: 650-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ ID NO: 931:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3152 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: circular
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: UNKNOWN
? ORIGINAL SOURCE:
? ORGANISM: PORPHYROMONAS GINGIVALIS
? FEATURE:
? NAME/KEY: misc.feature
? LOCATION: 1...3152
?
US-09-221-017B-931

Query Match
? Best Local Similarity 50.3%; Score 31.8; DB 4; Length 3152;
? Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 105 ATTAATTAATTCACAGGCTTTATCCGATGACTGGTTTGCACAAACAAAGATCCACT 164
DB 2914 AATGTTAAACCTATTTTGTGTGTGCGGATATCTTTGTTCCTCCCGAAGATCCCA 2973
QY 165 GACATTATAGGCTGCTTACAGCCTATACAGACATATGAAAGACTGTCAGGCTTATGA 224
DB 2974 GAGATCTTCCCATATGACAGCGCTTCCATATGACCTGCTGCAAAAGTAAAGAGATATC 3033
QY 225 ATTCACCTGGGAAGCTGAATGTGCAAAAGATTTT 259
DB 3034 CACAAACGGCTGATTTCTTGACACAAACATTTT 3068

RESULT 11
US-09-446-504-41
? Sequence 41 Application US/09446504
? Patent No. 6218150
? GENERAL INFORMATION:
? APPLICANT: UEMORI, Takashi
? APPLICANT: SATO, Yoshimi
? APPLICANT: FUJITA, Tomoko
? APPLICANT: MIYAKE, Kazuo
? APPLICANT: MUKAI, Hiroyuki
? APPLICANT: ASADA, Kiyozo
? APPLICANT: KATO, Ikunoshi
? TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
? FILE REFERENCE: 1422-408PCT
? CURRENT APPLICATION NUMBER: US/09/446.504
? CURRENT FILING DATE: 1999-12-23
? PRIOR APPLICATION NUMBER: PCT/JP98/02845
? PRIOR FILING DATE: 1998-06-24
? PRIOR APPLICATION NUMBER: JP 9-187496
? PRIOR FILING DATE: 1997-06-26
? PRIOR APPLICATION NUMBER: JP 9-320692
? PRIOR FILING DATE: 1997-11-27
? NUMBER OF SEQ ID NOS: 92
? SOFTWARE: Patent In Ver. 2.1
? SEQ ID NO 41
? LENGTH: 989
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
?
US-09-446-504-41

```


CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US

ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/238,163

FILING DATE: 03-MAY-1994

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 2307E-540

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2075 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 421..1401

US-08-238-163-3

Query Match 5.6%; Score 31.2; DB 1; Length 2075;

Best Local Similarity 45.5%; Pred. No. 2.9;

Matches 111; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 123 TTATCCGATGACGTGGTTTTCACAAACAAAGATACACGCTGACATTATGCTCTTA 182
DB 340 TATTTAGCAATACGGATTTTGGTCAGAAAGAACTGCCACAGCTACAGTACAG 281
QY 183 GAGCCTATCAAGCAATAGAGAGCTGCAAGCTTAATGCACTGCAAGTCTGAA 242
DB 280 GTGTATCTCAGTATCTAGATATTAATGCTGCTTTGAAATCAAACTCAAGAG 221
QY 243 ATGTCAAAAAGATTTAGCTCATGCAATTAAGACATTGATCTTTATACATAAAT 302
DB 220 ATGTCAATGATTTTACAAATTTATATCATGATGTTATATCATAGATATTTAAT 161
QY 303 GAGAAATCTGATTCGCAATACAAAGACCTTGCTTAATGAGAGTCCAAAGTTCT 362
DB 160 AATGATATTAATAAAGATACATAAAGTATTAATGTTTAAATGTAAAGTCTTT 101
QY 363 AGGT 366
DB 100 AAGT 97

RESULT 15

US-09-592-891A-13/C

Sequence 13: Application US/09592891A

Patent No. 6329174

GENERAL INFORMATION:

APPLICANT: Xiao-Zhou Michelle Wang

APPLICANT: Xavier Georges Sarda

APPLICANT: Michael David Tomalski

TITLE OF INVENTION: Helicobis Glutamate Receptor

FILE REFERENCE: A32815 07267 0118

CURRENT APPLICATION NUMBER: US/09/592,891A

CURRENT FILING DATE: 2001-01-22

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO: 13

LENGTH: 4621

TYPE: DNA

ORGANISM: Helicobis virescens

FEATURE:

NAME/KEY: unsure

LOCATION: (764)...(764)

OTHER INFORMATION: n is a, t, g, or c

US-09-592-891A-13

Query Match 5.6%; Score 31; DB 4; Length 4621;

Best Local Similarity 51.9%; Pred. No. 4.9;

Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 241 AAATGTCAAAAAGATTTTACCTCATGCAATTAAGACATTGATCTTTATACATAA 300
DB 2041 AAAATTAATCAAAAATTTTAAATCTCGAAAGTAAATATTTTGTACCAAAAT 1982
QY 301 TGGAGAAATCTGATTCGCAATTAATGAACCTTGCTTAATGAGAGTCCAAAGTT 360
DB 1981 TTGAGGACTTTATAGTGCAGATTAAGAGTTCTGCTTTAGATGACCAATCAAA 1922
QY 361 CTAGGCTCTTACCT 375
DB 1921 GTCTATTTTACCT 1907

Search completed: February 12, 2003, 15:12:59

Job time: 132 secs

QY 181 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAATG 240
DB 808 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAATG 867
QY 241 AAATGTGCAAAAAGATTTTACCTATGGAATTAAGCAATTCATTTATACCAATAAT 300
DB 866 AAATGTGCAAAAAGATTTTACCTATGGAATTAAGCAATTCATTTATACCAATAAT 927
QY 301 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 360
DB 928 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 987
QY 361 CTAGTCTTACCTGAG 420
DB 988 CTAGTCTTACCTGAG 1047
QY 421 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480
DB 1048 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1107
QY 481 ACCCAATGGGCGCTGGGCGTGTCTGTATTCATCATATGAGTGCATTTATCTGATTATC 540
DB 1108 TCCCAATGAG 1167
QY 541 AGTTCATGCGGCA 554
DB 1168 AGTTCATGCGGCA 1181

RESULT 2

US-09-734-017A-61
; Sequence 61, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehnhardt, Thomas
; APPLICANT: Knebel, Andreas
; APPLICANT: Clippus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duvenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09-734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin Ver. 2.1/WordPerfect
; SEQ ID NO 61
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(573)
; OTHER INFORMATION: 86_pprocl_094_g10
US-09-734-017A-61

Query Match 49.0%; Score 273.4; DB 10; Length 574;
Best Local Similarity 69.7%; Pred. No. 1.2e-69;
Matches 384; Conservative 0; Mismatches 166; Indels 1; Gaps 1;

QY 1 TGATGTGACCCCAATATTTATGATAGCATATATCTTCAATATTTGAGAGAGAGAG 60
DB 8 TCATATACCCAGAGTGTATATGATACGATATCTTTTGAATTTGTGATGATGATGTC 67

QY 61 GCCAAATGGAATACGTCGTCCTATTTGACCTGGAATTAAGCCATTAATTAATACAGG 120
DB 68 GTCAATATGATACCAAGAGGTCGTCATTTGACCTGTCATTCAGTCCCAATTAACAGG 127
QY 121 GCTTTATCCCATGATGAGGTTTTCGAAAACAAGATTAACAGCTGACATTTATGCTGCTT 180
DB 128 GCTTTATCCCATGATGAGGTCGTCATTTGACCTGTCATTCAGTCCCAATTAACAGG 187
QY 181 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAATG 240
DB 188 TAGAGCTATCAAGCAATGAGAGCTGTCAAGCTTATGGAATTCACCTGGGAATG 247
QY 241 AAATGTGCAAAAAGATTTTACCTATGGAATTAAGCAATTCATTTATACCAATAAT 300
DB 248 AAATGTGCAAAAAGATTTTACCTATGGAATTAAGCAATTCATTTATACCAATAAT 307
QY 301 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 360
DB 308 TGGAGAAATCGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 367
QY 361 CTAGTCTTACCTGAG 420
DB 368 CTAGTCTTACCTGAG 427
QY 421 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480
DB 428 CAATCTTTGGGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 487
QY 481 ACCCAATGGGCGCTGGGCGTGTCTGTATTCATCATATGAGTGCATTTATCTGATTATC 540
DB 488 TCCCAATGAG 546
QY 541 AGTTCATGCGG 551
DB 547 AGTTCATGCGG 557

RESULT 3

US-09-923-876-5596
; Sequence 5596, Application US/09923876
; Patent No. US2002013958A1
; GENERAL INFORMATION:
; APPLICANT: Laiguel, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 COM
; CURRENT APPLICATION NUMBER: US/09-923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5596
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457310H1
; LOCATION: 81..183
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5596

Query Match 22.3%; Score 124.4; DB 10; Length 264;
Best Local Similarity 75.5%; Pred. No. 1.2e-26;
Matches 166; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 340 TAATGAGAGTCGCAAGATTTCTAGTCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAG 399

```
Db 4 TAATTGAGGATGCCAAGGTTTCAAGGCATTTACCTTGAGGGCCAGCCAGCAATATGTTTTC 63
Qy 400 GTGTTAAAGAGATG-TCCGTCACATCTTTGGGCAATTCACCAAAACATCAATATCA 458
Db 64 GTGTTAAAGAGATGTTMCAGCATTATTTCTGGGCCAACAACCAAGAGACTATCTTAA 123
Qy 459 AGAGCAATGATGGATGATCAATCCAGCATGGGCGTGGGGGATTCCTGTTATCCATCA 518
Db 124 AGACATTTAGTTGGGATCAGTATCCCATGACCGTGGGGTGTGTTCTGGAAACCATCN 183
Qy 519 TATGGCATTTATCTGATTTATGATGATGCGGCCAGCTG 558
Db 184 CATGGAGCATTTACTGACCACTTCACATTCACAGACAGAG 223

RESULT 4
US-09-931-795-3
; Sequence 3, Application US/09931795
; Publication No. US20020198211A1
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/005003
; CURRENT APPLICATION NUMBER: US/09/931,795
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1983)
US-09-931-795-3

Query Match 20.9%; Score 116.4; DB 9; Length 2219;
Best Local Similarity 53.3%; Pred. No. 5.9e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

Qy 1 TGATGTGACCCCAATTTATATATGATATATCTCAAAATTTGTGAAGACTGTC 60
Db 683 TCATCATACGCGAGCTTTCTTGTGAGCTGACACATTTCTCCGCTTTGTGAAGAGATGA 742
Qy 61 GCCAAATTGGAAATACGTGTCTTATTTGACCTGGAAATTAAGCCATTAATTAATTAAG 120
Db 743 CCGACATGGGCAATCACTTGGCCCATGCTCCCGGGATTTCTTCCATTCACAGGCTTACACT 802
Qy 121 GCTTATCCGATGACTGGGTTTTCACAAACAAGATACAGCTGATATTAATGCTGCTT 180
Db 803 CCTTGGGAGCTTGTGAAGCTGTCCAAAGCTGGAGGATGACAGAGATCAAGAGCTGA 862
Qy 181 TAGAGCTTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGAATTAATTAATTAAGCTG 240
Db 863 TTGAGCAATCAAGAAACAAGATGCTGCATCCGAACTATGATGATGAGCTGAGCTGCTGA 922
Qy 241 AAATGTGCAAAAAGATTTTACCTGATGAATTAAGCAT--TGACATCTTATTAACCTAA 297
Db 923 GCTGTGTCAGAGAGCTTGTGGCAGTGGCTTGGTGCAGAGCTTCCACATCTTAACACCTCA 982
Qy 298 ATATGGAATATCTGATTTGGCATTAATTAAGAACTTGGCTTATTAAGAGATCCAAAG 357
```

```
Db 983 ACCGAGATGCTCTTACACACAGAGTCTGAAGCCGCGGGGAGATGTGACTGAGACCCA 1042
Qy 358 TTTCATAGTCTCTTACCTTGGAGACGCCCTGCAAAATGTTTCCGTGTTAAAGATGCTCC 417
Db 1043 GGGGTC--CCCTACCTGGGCTCTGATGAGCCACCCCAAGCCGCGAGAGAGATGTAC 1099
Qy 418 GTCCATCTTTTGGGCAAAATCGACCAAAAGCTACATATCAAGAGACATGAGAGGATG 477
Db 1100 GTCCATCTTTTGGGCGCTCCAGACCAAAAGATTAATCACTACCTACCCAGAGTGGAGC 1159
Qy 478 AATACCCATGAGCGCTTGGGATGATCTCTGAATTCATATATGATGATATCTGATT 537
Db 1160 AGTCCCTTAAGGCGCGCTGGGCAATTCCTCTCCCTTGGGAGCTGAAGAGCT 1219
Qy 538 ATCACTTCAT 547
Db 1220 ACTACCTCTT 1229

RESULT 5
US-09-728-910-3
; Sequence 3, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
; FILE REFERENCE: 04844/006001
; CURRENT APPLICATION NUMBER: US/09/728,910
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1983)
US-09-728-910-3

Query Match 20.9%; Score 116.4; DB 10; Length 2219;
Best Local Similarity 53.3%; Pred. No. 5.9e-24;
Matches 293; Conservative 0; Mismatches 251; Indels 6; Gaps 2;

Qy 1 TGATGTGACCCCAATTTATATATGATATATCTCAAAATTTGTGAAGACTGTC 60
Db 683 TCATCATACGCGAGCTTTCTTGTGAGCTGACACATTTCTCCGCTTTGTGAAGAGATGA 742
Qy 61 GCCAAATTGGAAATACGTGTCTTATTTGACCTGGAAATTAAGCCATTAATTAATTAAG 120
Db 743 CCGACATGGGCAATCACTTGGCCCATGCTCCCGGGATTTCTTCCATTCACAGGCTTACACT 802
Qy 121 GCTTATCCGATGACTGGGTTTTCACAAACAAGATACAGCTGATTAATTAATGCTGCTT 180
Db 803 CCTTGGGAGCTTGTGAAGCTGTCCAAAGCTGGAGGATGACAGAGATCAAGAGCTGA 862
Qy 181 TAGAGCTTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGAATTAATTAATTAAGCTG 240
Db 863 TTGAGCAATCAAGAAACAAGATGCTGCATCCGAACTATGATGATGAGCTGAGCTGCTGA 922
Qy 241 AAATGTGCAAAAAGATTTTACCTGATGAATTAAGCAT--TGACATCTTATTAACCTAA 297
Db 923 GCTGTGTCAGAGAGCTTGTGGCAGTGGCTTGGTGCAGAGCTTCCACATCTTAACACCTCA 982
Qy 298 ATATGGAATATCTGATTTGGCATTAATTAAGAACTTGGCTTATTAAGAGATCCAAAG 357
Db 983 ACCGAGATGCTCTACACAGAGTCTGAAGCCGCGGGGAGATGTGACTGAGACCCA 1042
```

Mon Feb 24 06:47:49 2003

us-09-720-451-5.rnpb

Page 4

| | | | |
|----|------|---|------|
| Qy | 358 | TTTTCAGGTCCTTACCTTGGAGAGCGCCGCAAAATGTTTCCGTGTTAAACAATATGCG | 417 |
| Db | 1043 | GGCGTC---CCCTACCGCTGGGCTCTCAGTGCCACCCCAAGGCCGAGAGAAATGTAC | 1099 |
| Qy | 418 | GTCCATCTTTTGGGCAATGACAAAAGGTATCATTTCAAGACACATAGATGATGTC | 477 |
| Db | 1100 | GTCCATCTTTTGGGCTTGGGCTCCACAGCAAAAGGTTNATNCTNACCTACCCGAGAGTGGACG | 1155 |
| Qy | 478 | AATACCAACATGGGGTGTGGGGTATCTCTGTAATACCAATATGTGATTAATCATGATT | 537 |
| Db | 1160 | AGTTTCCATACGGCCCGCTGGGCAATTCCTCTTCCCTTGTGGAGGCTGAAGAGACT | 1219 |
| Qy | 538 | ATCAGTTCAT | 547 |
| Db | 1220 | ACTACCTCTT | 1229 |

```

RESULT 6
US-09-931-795-1
Sequence 1, Application US/09931795
Publication No. US20020198211a1
GENERAL INFORMATION:
APPLICANT: ROZEN, Rita
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROPIRATE
FILE REFERENCE: 04844/005003
CURRENT APPLICATION NUMBER: US/09/931,795
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: US 09/592,595
PRIOR FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: US 09/258,928
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: US 08/738,000
PRIOR FILING DATE: 1997-02-12
PRIOR APPLICATION NUMBER: PCT/CA95/00314
PRIOR FILING DATE: 1995-05-25
PRIOR APPLICATION NUMBER: GB 9410620.0
PRIOR FILING DATE: 1994-05-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2220
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1980)
US-09-931-795-1

```

[illegible][illegible]

```

RESULT 7
US-09-728-910-1
Sequence 1, Application US/09728910
Patent No. US20010025030A1
GENERAL INFORMATION:
APPLICANT: Rozen, Rama
APPLICANT: Sekhon, Jaspreet
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE
FILE REFERENCE: 04844/006001
CURRENT APPLICATION NUMBER: US/09/728,910
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 09/258,928
PRIOR FILING DATE: 1999-03-01
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2220
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1980)
US-09-728-910-1

```

| Query Match | 20.9% | Score 116.4 | DB 10 | Length 2220 |
|-----------------------|----------------|--|----------|-------------|
| Best Local Similarity | 53.3% | Pred. No. 5.9e-24 | | |
| Matches 293 | Conservative 0 | Mismatches 251 | Indels 6 | Gaps 2 |
| Qy | 1 | TGATGTGACCCCAATTTATTTATGATGAGATATATCTCCAAATTTGTGAACGACATC | 60 | |
| Db | 683 | TCATCATACGACAGCTTTCTTGAGCGTGACACATTTCTCCGCTTTGTGAAGGATGCA | 742 | |
| Qy | 61 | GCCAAATTTGGATTAACGTGTCATATGTCACCGGATATATGCCCAATTAATTCGAAG | 120 | |
| Db | 743 | CCGACATGSGCATCAGTGGCCCATGTGTCCCGGATCTTCCCATCCAGGGCATACACT | 802 | |
| Qy | 121 | GCATTTCGCATGACGTGGTTTGTGAACCAAGATACAGCTGACATATAGCTGCTT | 180 | |
| Db | 803 | CCCTTGGCAGCTGTGTAACTGTCCAACTGTGAGAGTGCACAGGAGATCAAGAGCGCA | 862 | |
| Qy | 181 | TGAGGCTTATCAAGCAATGAAGAAGCTGTCAAGGCTTATGSAATTAACCTGGGAAC | 240 | |
| Db | 863 | TTTAGGCATTCAAAGACAAAGATGAGCCATTCGCACTATGSCATCGAGTGGCGCTGA | 922 | |
| Qy | 241 | AAATGTGCAAAAAGATTTTAGCTCATGAGATTTAAGACAT--TGACATCTTTATACATA | 297 | |
| Db | 923 | GCGTGTGCGAGACCTTCTGTGGCCAGTGGGTGGGCCAAGGCGTCCCATTTTACACCTTCA | 982 | |
| Qy | 298 | ATATGGAAGAAATCTGCATTGGCAATTAATGAACCTTGGCGCTAATATGAAAGTCCAAAG | 357 | |
| Db | 983 | ACGGGAGATGGCTATCCACAGAGAGTGTGAAGGCCCTGTGGGAGATGTGGACTAGAGGCCCA | 1042 | |

Page 5

```

1  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSER
2
3  FILE REFERENCE:  PL-0009 US
4
5  CURRENT APPLICATION NUMBER:  US/09/294, 093B
6
7  CURRENT FILING DATE:  1999-04-16
8
9  PRIOR APPLICATION NUMBER:  60/082,567
10
11  PRIOR FILING DATE:  April 21, 1998
12
13  NUMBER OF SEQ ID NOS:  6207
14
15  SOFTWARE:  PERL program
16
17  SEQ ID NO: 4653
18
19  LENGTH: 159
20
21  TYPE: DNA
22
23  ORGANISM:  Zea mays
24
25  FEATURE:
26
27  NAME/KEY:  misc_feature
28
29  OTHER INFORMATION:  Incyte ID NO.  US20010051335A1  700354/741H1
30
31  NAME/KEY:  unsure

```

```

: OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4653

Query Match      17.44; Score 97; DB 10; Length 159;
Best Local Similarity 75.66; Pred No. 7.9e-19;
Matches 118; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

```

| | | | |
|----|-----|---|------|
| QY | 388 | CAATGTTTTCOCGTGTAAGAAGATCTGCCTCAATCTTTGGCCAAATGACCAAAA | 44.7 |
| | | | |
| Db | 6 | CNATGTTTTCOCGTGTAAGAAGATCTGCACNATATTTTGGGCCAACACCAAAA | 65 |
| QY | 448 | GCTCATATCATGAGACCATGATGGATGATTAATACCCCATGAGGCCTTTGGGTATTCCT | 50.7 |
| | | | |
| Db | 66 | GCTATCTTAAAGACATTAATGTTGGATGATGATGCCCATGACGGTGGGATATCTC | 125 |
| QY | 508 | GPATCATCATATGATGTCATTTATCTGAATTACA | 54.1 |
| | | | |
| Db | 126 | GGAACCATCATATGAGCACTTATGACCAACA | 159 |

RESULT 10
 US-09-734-017A-63
 Sequence 63, Application US/09734017A
 Patent No. US20020142422A1
 GENERAL INFORMATION:
 APPLICANT: Ierchl, Jens
 APPLICANT: Benz, Andreas
 APPLICANT: Ehrhardt, Thomas
 APPLICANT: Reinold, Andreas
 APPLICANT: Cipurus, Petra
 APPLICANT: Bischoff, Friedrich
 APPLICANT: Frack, Markus
 APPLICANT: Duwend, Annette
 APPLICANT: Schmidt, Ralf-Michael
 TITLE OF INVENTION: Moss genes from *Physcomitrella patens* encoding proteins involved in
 TITLE OF INVENTION: The synthesis of amino acids, vitamins, cofactors, nucleotides and
 TITLE OF INVENTION: nucleosides
 FILE REFERENCE: EASF-NAS-1331-99-US
 CURRENT APPLICATION NUMBER: US/09/734,017A
 CURRENT FILING DATE: 2000-12-12
 PRIOR APPLICATION NUMBER: 60/171,100
 PRIOR FILING DATE: 1999-12-16
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: PatentIn Ver. 2.1/WordPerfect
 SEQ ID NO 63
 LENGTH: 409
 TYPE: DNA
 ORGANISM: *Physcomitrella patens*
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (2)..(409)
 OTHER INFORMATION: 62_rmm20_c10rev
 US-09-734-017A-63

Mon Feb 24 06:47:49 2003

us-09-720-451-5.rnpb

Page 6

```

Query Match Similarity      15.3%; Score 85.6; DB 10; Length 409;
Best Local Similarity       Pred. No. 2,de-15;
Matches 114; Conservative   0; Mismatches 29; Indels 1; Gaps 1;

OY      1 TGAATGTCACCCAAATTATTATTTANGATAGCATATATCTCACAATTGTGACAGACTGC 60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      261 TCATATTACACCCECGCTGTTTATGATATACCGCANACTCTTTTGAAATTTGTGAAAGATATGTC 320

OY      61 GCCAAAATGGSAATAACGTCGTCTATTGTACTCCGGAAATTATGCCATTATATTAAGA-G 119
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      321 GTCAAAATGGATACAAGGTGCCCTTGTACTCGTATCATAGCCCAATCAAATTAACAAG 380

OY      120 GGCTTATPCCGATGACTGGGTTT 143
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      381 GGCTTTCTCCGATGACCACTTT 404


RESULT 11
US-10-113-852A-1
: Sequence 1, Application US/10113852A
: Publication NO. US20020192784A1
: GENERAL INFORMATION:
: APPLICANT: Appling, Dean R.
: APPLICANT: Hanson, Andrew D.
: APPLICANT: Raymond, Rhonda K.
: TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast Strain
: FILE REFERENCE: 119921-1053
: CURRENT FILING DATE: 2002-03-29
: PRIOR APPLICATION NUMBER: 60/280,333
: PRIORITY FILING DATE: 2001-03-30
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 3082
: TYPE: DNA
: ORGANISM: Saccharomyces cerevisiae
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (826)..(2625)
: OTHER INFORMATION: DNA construct
: US-10-113-852A-1
```

| | Query Match | Best Local Match | Similarity | Conservative | Score | DB | Length | Indels | Caps |
|--------------|-------------|---|------------|--------------|--------------------|-------|--------|--------|------|
| Matches 274: | 12.0%; | 12.0%; | 49.58; | 0; | 67; | 3082; | | | |
| | | | | | Pred. No. 1.3e-09; | | | | |
| | | | | | Mismatches 255; | | | | |
| | | | | | | | | | |
| QY | 1 | TGATGTGCACCAAAATTTATTTATGATGACGATATATATCTCCAAATTTGTAAGACATGTC | 60 | | | | | | |
| Db | 1376 | TCATCATCATCATCGATGGTTTATAGAGTGTGATATTTACTACAAACGGGTGCTCCAAAGTTA | 1435 | | | | | | |
| QY | 61 | GCCAAATGTGAATACGTGCTGTCATATGTGACCTGGAATTTAGCCATATATAATTAATGAAG | 120 | | | | | | |
| Db | 1436 | GAGGTGGGGGACATGACGTGCCATATATTCGGGGATCATGCCATCTACTACAGGGG | 1495 | | | | | | |
| QY | 121 | GCATTATCCGACATGACTGGGTTTGCAAAACAAAGATACACAGCTATATAGCGCGCTT | 180 | | | | | | |
| Db | 1496 | CCCTCTGGAAGAAGATCCAAATGGGGCCAAATCTCATCCCTGCATATCTTCGTCCGCTT | 1555 | | | | | | |
| QY | 181 | TAGAGCCTATCAAGACATGAAGAAGAGCGTCAAGAGGCTATGAGATTCACCTGGGAATG | 240 | | | | | | |
| Db | 1556 | TGGAATCTATCAAGGACAGCATGAGAGATTGGTCTGATATGCGAACTACTTATATGTCGG | 1615 | | | | | | |
| QY | 241 | AAATGTGCAAAAAGATTTTATAGCGATG---AAATGAAGCATGATCTTATATACACGTA | 297 | | | | | | |
| Db | 1616 | AAATGTCTCAAAAATTTGCTGACAGAGTGTTGACGTTTCTACTTGACATCTCAACACATGA | 1675 | | | | | | |
| QY | 299 | ATATGGAGAAATCTGCATTTGGCAATACATATGAACCTTGCGCTATTTGAAGAGTCCAAAG | 357 | | | | | | |
| Db | 1676 | ACTTGGAAAAAGGCGCTCTCATGATATCTTGGAAATATGAACATCTCTACAGGAGATG | 1735 | | | | | | |
| QY | 358 | TTTCTAGGTC-----CTTACCTTGGAGACGCCCTGCACAAATGTTTTCCGTG | 402 | | | | | | |

| | | | |
|----|------|---|------|
| Db | 1736 | AGTTTATATGACATCTCCATTTGGCCCTGTGCCATGAGAAAAATCTTTGATCTCAAAACGTA | 1739 |
| | | | |
| QY | 403 | TTAAAGAAAGATGTGCGCTCCATCTCTTTTGGGCAAAATGCACAAAAAGCTACATTTCAAGGA | 462 |
| | | | |
| Db | 1796 | AAAAAGAGGAATGACACCTATCTCTGGAGAGAAAGACCTTCTCTCATATGTGGCAAGAA | 1855 |
| | | | |
| QY | 463 | CCATAGAGATG-----GGATCAATATCCCATCATGAGCGCTTGGGGTGATTTCTGTATTCAT | 516 |
| | | | |
| Db | 1856 | CCCTTCATATGGGCCGTGGAGCAATATCCCAACAGGAGATTTGGTGATTTGCTTCTCTCG | 1915 |
| | | | |
| QY | 517 | CATATGTGTCATT | 529 |
| | | | |
| Db | 1916 | CTTTCGGTAGCTTT | 1928 |
| | | | |

```

RESULT 12
US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. US20020127687A1
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: MATANABE, HIDEMI
APPLICANT: HATTORI, MASAHISA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

```

| Query Match | 9.6% | Score 53.4 | DB 10 | Length 640681 |
|-----------------------|----------------|---|----------|---------------|
| Best local Similarity | 47.5% | Pred. No. 0.00011 | | |
| Matches 159 | Conservative 0 | Mismatches 176 | Indels 0 | Gaps 0 |
| Oy | 7 | TACACCAATATTTATGATGCGATATATTCCTCAAAATTTGTGTAAGCACTGCGCCAA | 66 | |
| Db | 50707 | TTACACAGTCTTTTATATATGGAAGTATTTTCGCTTTGCGAGATATGATATATAA | 50766 | |
| Oy | 67 | TTGGAAATACAGTGTCTCATGTCACGTAATTAATGCGCATTAATATACAGGCTTA | 126 | |
| Db | 50767 | ATTAATAAT | 50826 | |
| Oy | 127 | TCCGATGACTGGGTTTGCCAAAACAAAGATACAGCTGACACTATGCGCTTTAGAC | 186 | |
| Db | 50827 | AACGTTTTTCAAGTATGACTAATGTTAAATATCAAAATGATGCTGATATGTTTAATG | 50886 | |
| Oy | 187 | CTATCAAGGACCAATGAGAGAGCTGTCAAGGCTTATGGAATTCACCTGGAGCTGAATGT | 246 | |
| Db | 50887 | GTTTAGTATGATATGATATATTCACAGAAAATTAATATGCTGTATGCTATGATATGG | 50946 | |
| Oy | 247 | GCAAAAGATTTTACCTCATGAGATTAAGACATTTGACTTTTATACATAATATATGAGA | 306 | |
| Db | 50947 | TAAATAAATATCTGTGAGAGAGGTAAAAAATTTTCAATTTTACACTTAAATCAGTGTG | 51006 | |
| Oy | 307 | AATTCGATTCGCAATPACTAATGACCTGGAGCTA | 341 | |
| Db | 51007 | ATATTACTTATTTCTATTGTCATATCTTAAGTCTCA | 51041 | |

RESULT 13
US-09-75d-853a-1/c
; Sequence 1, Application US/09754853a
; Publication No. US20030005491A1
; GENERAL INFORMATION:

```

; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 1
; LENGTH: 127197
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 515002_region_g2
US-09-754-853A-1

```

```

Query Match
Best Local Similarity 55.98; Score 40; DB 9; Length 127197;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

```

QY 193 AGACATGAGAGAGCTGCAAGCTTATGGAATTCACCTGGAGCACTGAAATGTGCAAAA 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58616 AGGGGAATGATGAGAGGTCAGCCCTAAGAAAAGATTTCACCTTGTCAAAAAGGGGCGAGA 58557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 253 AGATTTTGGCTGATGAGATTAAGACATGTCATCTTATACCTTAATATGAGAAATCTG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58556 AGAAGGCTCAAAATGATTTGAGATCTCATCATGATACCTTTATCAAGAGGTACAAAG 58497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 313 CATGCGCAATCTAAT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58496 CATCTGGAACAGAAAT 58481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-09-815-242-7047
; Sequence 7047, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7047
; LENGTH: 1266

```

```

; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1266)
US-09-815-242-7047

```

```

Query Match
Best Local Similarity 60.64; Score 34.8; DB 10; Length 1266;
Matches 57; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

QY 353 CAAAGTTCTAGGTCCTTACTTGGAGAGCCCTGCAATGTTTCCGTGAAGAGA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 CAAATTTATTAACCTCGTACTTATGGAATTAAGTGGGATGCTTAATGATTAAGAGA 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 413 TGTGCTCAATCTTTGGCAATGACCAAAA 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 TGTGCTCAAAAAGCATTAAGATGTAACCAAAA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-938-842A-4584/C
; Sequence 4584, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCHRIP300-3
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4584
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4584

```

```

Query Match
Best Local Similarity 49.2%; Score 33.4; DB 9; Length 2000;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```

```

QY 273 AAGACATTCATCTTATACCTTAATATGAGAAATCGCTTGGCAATCTAATGAC 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 AAGACATTAACATTTTAAAGTTAGATTAAGTTTGAAGTGAAGAAAAAATCAAA 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 333 CTGGCGTAATGAGAGCTCAAAAGTTCTAGTCTTACCTTGGAGAGCCCTGCAAT 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 ATTATTTTAAATCATTTATCCAAATCATTAATCCAAACGATGCCACCTTAAT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 GTTTCGGTGAAGAGATGTCGCAATCTTTGGCAATGACCAAAAAGCTA 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 635 ATTAACTATTTAATAAGTTTGAAGCATAGCTATATATAAAAGCTAAAAAATATA 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: February 23, 2003, 19:39:34
Job time : 240 secs

7
.
.
.

Mon Feb 24 06:47:51 2003

us-09-720-451-5.rst

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2003, 13:48:04 : Search time 2194 Seconds
(without alignments)
4118.998 Million cell updates/sec

Title: US-09-720-451-5
Perfect score: 558
Sequence: 1 tgaattaccacattatc.....tcaagtcacgcgcacg 558

Scoring table: IDENTITY-NUC
Gapex 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estov:*
6: em_estpl:*
7: em_estpl:*
8: em_estpl:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 556.4 | 99.7 | 588 | 12 | BF071278 |
| 2 | 523.6 | 93.8 | 565 | 9 | AT900449 |
| 3 | 484.4 | 86.8 | 759 | 13 | BT308605 |
| 4 | 482.8 | 86.5 | 755 | 13 | BT308302 |
| 5 | 466.4 | 83.6 | 573 | 10 | AM203249 |
| 6 | 466.4 | 83.6 | 579 | 10 | AM203207 |

| | | | | | | | |
|----|-------|------|------|----|-----------|-----------|------------|
| 7 | 429.4 | 77.0 | 468 | 12 | BG405762 | BG405762 | sac37108. |
| 8 | 422 | 75.6 | 668 | 14 | BQ404086 | BQ404086 | GA_Ed006 |
| 9 | 400.4 | 71.8 | 589 | 12 | BF269445 | BF269445 | GA_ED000 |
| 10 | 399.2 | 71.5 | 593 | 10 | BE204232 | BE204232 | EST396908 |
| 11 | 378 | 67.7 | 558 | 10 | BE433609 | BE433609 | EST400138 |
| 12 | 376.6 | 67.5 | 556 | 10 | BE433601 | BE433601 | EST400130 |
| 13 | 369.2 | 66.2 | 661 | 14 | BQ863196 | BQ863196 | GC23015. |
| 14 | 366.8 | 66.7 | 652 | 12 | BG904512 | BG904512 | Tair1132G |
| 15 | 364.6 | 65.3 | 560 | 13 | BM526290 | BM526290 | sal19404. |
| 16 | 362.2 | 64.9 | 592 | 10 | BE354491 | BE354491 | EST35834 |
| 17 | 358 | 64.2 | 2423 | 11 | AT109096 | AT109096 | Zea_mays |
| 18 | 357.6 | 64.1 | 569 | 14 | BM855389 | BM855389 | sal199b11. |
| 19 | 351 | 62.9 | 657 | 14 | BQ471589 | BQ471589 | HV020031 |
| 20 | 348.8 | 62.5 | 419 | 10 | AM234903 | AM234903 | sif20e05.Y |
| 21 | 344.6 | 61.8 | 568 | 9 | AT729106 | AT729106 | BHLGH1126 |
| 22 | 338.6 | 60.7 | 505 | 10 | AM029711 | AM029711 | EST72966 |
| 23 | 338.6 | 60.5 | 516 | 14 | BQ446637 | BQ446637 | Tae15007F |
| 24 | 337.8 | 60.5 | 582 | 10 | AM649374 | AM649374 | EST327828 |
| 25 | 335 | 60.0 | 586 | 10 | AM931213 | AM931213 | EST357056 |
| 26 | 333 | 59.7 | 397 | 12 | BR070260 | BR070260 | sl14b03.Y |
| 27 | 332.2 | 59.5 | 523 | 10 | BE403916 | BE403916 | WHE0415_G |
| 28 | 329 | 59.0 | 530 | 13 | B1128372 | B1128372 | IO63P12P |
| 29 | 328.2 | 58.8 | 627 | 10 | AV915401 | AV915401 | AV915401 |
| 30 | 326 | 58.4 | 595 | 14 | BQ471487 | BQ471487 | HV02J07F |
| 31 | 322.6 | 57.8 | 678 | 9 | AT727725 | AT727725 | BHLGH1883 |
| 32 | 317.4 | 56.9 | 685 | 13 | BE1921517 | BE1921517 | EST541420 |
| 33 | 317 | 56.8 | 538 | 10 | BE346411 | BE346411 | sp25c05.Y |
| 34 | 317 | 56.8 | 634 | 10 | BE346389 | BE346389 | sp25a05.Y |
| 35 | 313.8 | 55.2 | 538 | 13 | B1131003 | B1131003 | G113P82Y |
| 36 | 311.8 | 55.9 | 642 | 14 | BQ862545 | BQ862545 | OGC21606Y |
| 37 | 311.4 | 55.8 | 814 | 10 | BE194793 | BE194793 | HVSMEN008 |
| 38 | 308.2 | 55.2 | 438 | 13 | B1129696 | B1129696 | C097P07Y |
| 39 | 305 | 54.7 | 760 | 14 | BQ924876 | BQ924876 | OGRE621.Y |
| 40 | 302.6 | 54.2 | 548 | 13 | B1921139 | B1921139 | EST194104Z |
| 41 | 289.4 | 51.9 | 511 | 13 | B1131121 | B1131121 | G115P70Y |
| 42 | 286.6 | 51.6 | 613 | 14 | BQ402630 | BQ402630 | GA_Ed005 |
| 43 | 286.6 | 51.6 | 535 | 10 | BE215671 | BE215671 | HV_CED000 |
| 44 | 283 | 50.7 | 387 | 13 | B1127775 | B1127775 | G065P81Y |
| 45 | 281.6 | 50.5 | 700 | 10 | BE576981 | BE576981 | LO-1588T3 |

ALIGNMENTS

| | | | | | |
|---|----------|--------|------|--------|-----------------|
| RESULT 1 | BF071278 | 588 bp | mRNA | linear | EST 06-DEC-2001 |
| LOCUS | BF071278 | | | | |
| DEFINITION | BF071278 | 588 bp | mRNA | linear | EST 06-DEC-2001 |
| ACCESSION | BF071278 | | | | |
| VERSION | BF071278 | | | | |
| KEYWORDS | BF071278 | | | | |
| SOURCE | BF071278 | | | | |
| ORGANISM | BF071278 | | | | |
| REFERENCE | BF071278 | | | | |
| AUTHORS | BF071278 | | | | |
| Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wille, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R. | | | | | |
| Public soybean EST Project | | | | | |
| Unpublished (1999) | | | | | |
| Contact: Shoemaker R/Public Soybean EST Project | | | | | |
| Public Soybean EST Project | | | | | |
| Washington University School of Medicine | | | | | |
| 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA | | | | | |
| Tel: 314 286 1800 | | | | | |

Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: curesgen.com
 Insert length: 1779 Std Error: 0.00
 High quality sequence stop: 414.

FEATURES

source

1. 588
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl012-1734"
 /clone_1lb="Gm-cl067"
 /tissue_type="germinating shoot, 3 day old seedling, auxin
 treatment"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from germinating shoots of 3 day old seedling for the
 cultivar Williams 82. The seedlings were germinated in a
 growth chamber using germination paper in a solution
 containing 100ppm auxin. Complementary DNA was synthesized
 from mRNA using a primer consisting of a poly(dT) sequence
 with a XhoI restriction site. EcoRI adapters were ligated
 to the blunt-ended cDNA fragments followed by XhoI
 digestion. The cDNA fragments were directionally cloned
 into the EcoRI-XhoI restriction site of the pBluescript
 vector. The ligated cDNA fragments were transformed into
 DH10B host cells (GibcoBRL). This library was constructed
 in the laboratory of Dr. Randy Shoemaker."

BASE COUNT

182 a 114 c 124 g 168 t

ORIGIN

Query Match 99.7% Score 556.4; DB 12; Length 588;
 Best Local Similarity 99.8% Pred. No. 6,6e-153;
 Matches 557; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGATGTCACCCATTTTATGATGAGATATTCCTCAATTTTGACAGCTGTC 60
 DB 5 TGATGTCACCCATTTTATGATGAGATATTCCTCAATTTTGACAGCTGTC 64
 QY 61 GCCAATGGAATACGCTGCTATGACCTGGAATATGATTAATTAACAG 120
 DB 65 GCCAATGGAATACGCTGCTATGACCTGGAATATGATTAATTAACAG 124
 QY 121 GCTTATCCGATGCTGCTTATGCAAAACAAAGATACGCTGATATGCGCTT 180
 DB 125 GCTTATCCGATGCTGCTTATGCAAAACAAAGATACGCTGATATGCGCTT 184
 QY 181 TAGAGCCATCAAGACATGAAGAGCTGTCAGGCTTATGCAATTCACCTGGACCTG 240
 DB 185 TAGAGCCATCAAGACATGAAGAGCTGTCAGGCTTATGCAATTCACCTGGACCTG 244
 QY 241 AATATGCAAAAGATTTAGCTATGATGAATTAAGCATTCATCTTATACATAAATA 300
 DB 245 AATATGCAAAAGATTTAGCTATGATGAATTAAGCATTCATCTTATACATAAATA 304
 QY 301 TCGAAGAAATCTGCTTATGCAATTAAGCAATTCGCTGATTAAGAGAGTCCAAAGTTT 360
 DB 305 TCGAAGAAATCTGCTTATGCAATTAAGCAATTCGCTGATTAAGAGAGTCCAAAGTTT 364
 QY 361 CTAGTCTTACCTTGGAGACGCTCGCAATGTTTCCGTGTTAAAGAAATGTCGCTC 420
 DB 365 CTAGTCTTACCTTGGAGACGCTCGCAATGTTTCCGTGTTAAAGAAATGTCGCTC 424
 QY 421 CATCTTTTGGCAAAATCGACAAAAGCTACATCAAGAGCATTAGATGGAGATCAAT 480
 DB 425 CATCTTTTGGCAAAATCGACAAAAGCTACATCAAGAGCATTAGATGGAGATCAAT 484
 QY 481 ACCCAATGGGCGTTGGGGTGAATCTGTAATCAATATGCGCAATTAATCGATTAAT 540
 DB 485 ACCCAATGGGCGTTGGGGTGAATCTGTAATCAATATGCGCAATTAATCGATTAAT 544

QY 541 AGTTCATGCGCGCACGCTG 558
 DB 545 AGTTCATGCGCGCACGCTG 562

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 565)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khama

A., Bolla,B., Merris,M., Hillier,L., Kueba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Rheising,B., Allen,M., Boxers

J., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck

R., Ritter,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., Mccann

R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Possible reversed clone: similarity on wrong strand This clone is

available through: ResGen, Invitrogen Corp. 2130 South Memorial

Parkway Huntsville, AL 35801 For further information call: (800

)-533-4363 or contact via email: curesgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 401.

Location/Qualifiers

1. 565

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl012-1441"

/clone_1lb="Gm-cl012"

/tissue_type="Apical shoot tips, 9-10 day old etiolated

seedlings"

/lab_host="XL10-Gold"

/note="Vector: pBluescript II XE; Site_1: EcoRI; Site_2:

XhoI; This cDNA library was constructed from mRNA isolated

from the apical shoots of 9 to 10 day old etiolated

seedlings. The shoot tips including any emerged leaves

were harvested for mRNA isolation. The cDNA library was

prepared using the Stratagene pBluescript II XE cDNA

library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were

transformed into XL10-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Erpelting."

BASE COUNT 171 a 112 c 117 g 163 t 2 others

ORIGIN

Query Match 93.8% Score 523.6; DB 9; Length 565;
 Best Local Similarity 97.3% Pred. No. 2.9e-143;
 Matches 532; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

OY 12 CAATATTTTATGATAGCATATATTTCTCAATTTTGTAAGCACTGCGCCAAATTGGA 71
    |||||||
Db 2 CAGTATATTTTATGATAGCATATATTTCTCAATTTTGTAAGCACTGCGCCAAATTGGA 61
OY 72 ATAGCTGTCCTATTTGACCTGGATTTATGCCATTTAATTAATACAAAGGCTTTATCCGC 131
    |||||||
Db 62 ATAGCTGTCCTATTTGACCTGGATTTATGCCATTTAATTAATACAAAGGCTTTATCCGC 121
OY 132 ATAGCTGTCCTATTTGACCTGGATTTATGCCATTTAATTAATACAAAGGCTTTATCCGC 191
    |||||||
Db 122 ATAGCTGTCCTATTTGACCTGGATTTATGCCATTTAATTAATACAAAGGCTTTATCCGC 181
OY 192 AAGGACATGAGAAGAGCTGTCAAGGCTTTGGAATTTCCCTGGAGACTGAAATGTCGAA 251
    |||||||
Db 182 AAGGACATGAGAAGAGCTGTCAAGGCTTTGGAATTTCCCTGGAGACTGAAATGTCGAA 241
OY 252 AAGATTTAGCTCATGGAATTTAAGACATTTGCAATTTTATACCTAATATGAGAAATCT 311
    |||||||
Db 242 AAGATTTAGCTCATGGAATTTAAGACATTTGCAATTTTATACCTAATATGAGAAATCG 301
OY 312 GCATTTGGCAATTAATGAACCTTGGCTAATTTGAAGAGTCAAAAGTTTCTAGTCTTTA 371
    |||||||
Db 302 GCATTTGGCAATTAATGAACCTTGGCTAATTTGAAGAGTCAAAAGTTTCTAGTCTTTA 361
OY 372 CCTTGAGAGCGCCCTGCAAAATGTTTCCGTTTAAGAAGATGTCGCTCAATCTTTGG 431
    |||||||
Db 362 CCTTGAGAGCGCCCTGCAAAATGTTTCCGTTTAAGAAGATGTCGCTCAATCTTTGG 421
OY 432 GCAATTCGACGCAAAAGCTCATATGCAAGACATGAGATGAGATGCAATACCAATGG 491
    |||||||
Db 422 GCAATTCGACGCAAAAGCTCATATGCAAGACATGAGATGAGATGCAATACCAATGG 481
OY 492 CGTTGGGGTGAATTCCTGTAATCATCATATGCTGATTAATGATTTACGTTTCATGGG 551
    |||||||
Db 482 CGTTGGGGTGAATTCCTGTAATCATCATATGCTGATTAATGATTTACGTTTCATGGG 541
OY 552 CCACGATG 558
    |||||||
Db 542 TCACGTG 548

RESULT 3
LOCUS B1308605 759 bp mRNA linear EST 20-JUL-2001
DEFINITION EST30015 GP0D Medicago truncatula cDNA clone pGP0D-714 5' end,
ACCESSION B1308605
VERSION B1308605
KEYWORDS mRNA sequence.
SOURCE Medicago
ORGANISM Medicago truncatula
            battef medic.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
            Medicago.
            1 (bases 1 to 759)
REFERENCE Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Uterback,T., Cho
AUTHORS J., and Fraser,C.M.
TITLE ESTs from developing reproductive tissues of Medicago truncatula
JOURNAL Unpublished (2001)
COMMENT Contact: Michael A. Grusak
          USDA/ARS Children's Nutrition Research Center
          Baylor College of Medicine
          1100 Bates street, Houston, TX 77030-2600, USA
          Tel: 713-798-7044
          Fax: 713-798-7078
          Email: mgrusak@bcm.tmc.edu
          B593460e
          TIGR sequence name: MTOAR50TK
          More information is available at: www.medicago.org
          Seq primer: SKmod (CTA gaa cta gta gat cc).
          Location/Qualifiers

```

```

source
1. 759
/organism="Medicago truncatula"
/cultivar="Al7"
/db_xref="taxon:3880"
/clone="pGP0D-714"
/tissue_type="immature pod walls"
/seq_stage="immature pods, ranging in age from 15 to 30
days after pollination"
/notes="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Immature pods, ranging in age from 15 to 30 days
after pollination, were collected from greenhouse-grown
plants. At harvest, seeds were removed from pods and
isolated pod walls were collected and immediately frozen
in liquid nitrogen. Pod walls were pooled for mRNA
extraction. cDNA was prepared from polyA-enriched RNA.
The cDNA was directionally ligated into the NotIap XR
vector from Stratagene and packaged using GigaPack III
Gold packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-Zap phage using
Ex-assist Helper phage and propagated in XL10R cells."
BASE COUNT 235 a 143 c 175 g 206 t
ORIGIN
Query Match 86.8%; Score 484.4; DB 13; Length 759;
Best local similarity 91.8%; Pred. No. 1.2e-131;
Matches 512; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

OY 1 TGATTTGACCCCAATTTTATGATAGCATATATTTCTCAATTTTGGAACGACTGTC 60
    |||||||
Db 40 TGATTTGACCCCAATTTTATGATAGCATATATTTCTCAATTTTGGAATGATGTC 99
OY 61 GCCAATTTGGAATAGCTGTCTATGCTACGTAATTTGCCATTAATTTATTTCAAG 120
    |||||||
Db 100 GCCAATTTGGAATAGCTGTCTATGCTACGTAATTTGCCATTAATTTCAAG 159
OY 121 GCTTTATCCGATGACGTGGGTTTGCAGAAACAAAGATACAGCTGACATTAAGCTGTT 180
    |||||||
Db 160 GCTTTATCCGATGACGTGGGTTTGCAGAAACAAAGATACAGCTGACATTAAGCTGTT 219
OY 181 TAGAGCTATGACGACAAATGAAGAGCTGTCAAGCTTTGGAATTCACCTGGACGTG 240
    |||||||
Db 220 TAGAGCTATGACGACAAATGAAGAGCTGTCAAGCTTTGGAATTCACCTGGACGTG 279
OY 241 AATATGCAAAAGATTTAGCTCATGGAATTAAGACATTTGCAATCTTTATACCTAATA 300
    |||||||
Db 280 AATATGCAAAAGATTTAGCTCATGGAATTAAGACATTTGCAATCTTTACAGCTAATA 339
OY 301 TGGAGAAATCGCATTTGGCAATTAATGACCTTGGCTTAATGAAGAGTTC 360
    |||||||
Db 340 TGGAGAAATCGCATTTGGCAATTAATGACCTTGGCTTAAATGAAGAGTTC 399
OY 361 CTAGGTCCTTACCTTTGGAGAGCGCCCTGCAAAAGCTTTCCGTTAAAGAAATGTCGCTC 420
    |||||||
Db 400 CTAGGTCCTTACCTTTGGAGAGCGCCCTGCAAAAGCTTTCCGTTAAAGAAATGTCGCTC 459
OY 421 CAATCTTTTGGCAATTTGACCAAAAAGCTTACATTAAGACCATTAAGATGGATCAAT 480
    |||||||
Db 460 CAATCTTTTGGCAATTTGACCAAAAAGCTTACATTAAGACCATTAAGATGGATCAAT 519
OY 481 ACCCAATGAGGCGTTGGGATTTCTGTATATCATATATGATGATTAATGATATC 540
    |||||||
Db 520 ACCCAATGAGGCGTTGGGATTTCTGTATATCATATATGATGATTAATGATATC 579
OY 541 AGTTCAATGGCGGACGATG 558
    |||||||
Db 580 AGTTCAATGGCGGACGATG 597

RESULT 4
LOCUS B1308302 755 bp mRNA linear EST 20-JUL-2001
DEFINITION EST529712 GP0D Medicago truncatula cDNA clone pGP0D-2P15 5' end,

```


[illegible]

| FEATURES | | Location/Qualifiers | |
|---------------------------|--|--|--|
| source | | 1..468 | |
| | | /organism="Glycine max" | |
| | | /db_xref="taxon:3847" | |
| | | /clone="GENOME SYSTEMS CLONE ID: Gm-c1051-4503" | |
| | | /clone_1lb="Gm-c1051" | |
| | | /tissue_type="Floral meristematic mRNA" | |
| | | /lab_host="DH10B" | |
| | | /note="Vector: plusescript II SK+; site_1: EcoRI; site_2: XhoI. The cDNA library was constructed from floral meristematic mRNA provided by Dr. Haila Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The EcoRI-XhoI restriction site of the plusescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker." | |
| BASE COUNT | 146 a 87 c 96 g 139 t | | |
| ORIGIN | | | |
| Query Match | 77.0%; Score 429.4; DB 12; Length 468; | | |
| Best Local Similarity | 95.5%; Pred. No. 1.5e-115; | | |
| Matches 442; Conservative | 0; Mismatches 21; Indels 0; Gaps 0; | | |
| Oy | 1 TGAATGTCACCCAAATATTTATGATAGAGATATATCTCCTCAATTTGGAACAGCTGC 60 | | |
| Db | 5 TGAATGTCACGATATTTATGATAGAGATATATCTCCTCAATTTGGAACAGCTGC 64 | | |
| Oy | 61 GCCAAATTTGGAATPAGCTGTCATTTGACCGGGAAATTTAGCCATTTAATTAACAAG 120 | | |
| Db | 65 GCCAAATTTGGAATPAGCTGTCATTTGACCGGGAAATTTAGCCATTTAATTAACAAG 124 | | |
| Oy | 121 GCTTTATCCGATGATCGGTTTGGCAAAAGAAATPACAGCTGACATTAATGGCTCTT 180 | | |
| Db | 125 GCTTTATCCGATGATCGGTTTGGCAAAAGAAATPACAGCTGACATTAATGGCTCTT 184 | | |
| Oy | 181 TAGAGCTATCAAGACAAATGAAGAGCTGACGCTTATGGAATTAACCGGGAACTG 240 | | |
| Db | 185 TAGAGCTATCAAGACAAATGAAGAGCTGACGCTTATGGAATTAACCGGGAACTG 244 | | |
| Oy | 241 AATATGCAAAAAGATTTTACGTCATGATTAAGACATTCATCTTTATACATAATA 300 | | |
| Db | 245 AATATGCAAAAAGATTTTACGTCATGATTAAGACATTCATCTTTATACATAATA 304 | | |
| Oy | 301 TGGAGAAATGCAATGGCAATACATTAAGACATTCGCTTAATGAGAGTCCAAAGTTT 360 | | |
| Db | 305 TGGAGAAATGCAATGGCAATACATTAAGACATTCGCTTAATGAGAGTCCAAAGTTT 364 | | |
| Oy | 361 CTAGAGCTCTTACCTTGGAGAGCGCCCTGCAAAATGTTTCCGCTTTAAAGAGATGTCGC 420 | | |
| Db | 365 CTAGAGCTCTTACCTTGGAGAGCGCCCTGCAAAATGTTTCCGCTTTAAAGAGATGTCGC 424 | | |
| Oy | 421 CAATCTTTTGGGCAAAATGACCAAAAGCTACATTAAGAGAC 463 | | |
| Db | 425 CAATCTTTTGGGCAATTTGCTCAAAAGCTCATTAATTAAGAGAC 467 | | |
| RESULT 8 | | | |
| LOCUS | BO404086 668 bp mRNA linear EST 22-MAY-2002 | | |
| DEFINITION | GA_E00066A09f Gossypium arboreum 7-10 dpa fiber library Gossypium | | |
| ACCESSION | BO404086 | | |
| VERSION | BO404086.1 GI:21091773 | | |
| KEYWORDS | EST; | | |
| SOURCE | Gossypium arboreum. | | |
| ORGANISM | Gossypium arboreum. | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Malvales; Malvaceae; Gossypium. | | |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|-----------|---|--|---------------------|--------------------|
| 1 | (bases 1 to 668) | Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry | | |
| 2 | D., Wood, T. C., Leslie, A. and Wilkins, T. A. | An integrated analysis of the genetics, development, and evolution | | |
| 3 | of the cotton fiber | Unpublished (2000) | | |
| 4 | Contact: Wing, RA | | | |
| 5 | Clemson University Genomics Institute | | | |
| 6 | Clemson University | | | |
| 7 | 100 Jordan Hall, Clemson, SC 29634, USA | | | |
| 8 | Tel: 864 656 7288 | | | |
| 9 | Fax: 864 656 4293 | | | |
| 10 | Email: twing@clemson.edu | | | |
| 11 | Total High Quality Bases = 504 | | | |
| 12 | Seq primer: TAATACGCTCACTTAAGG | | | |
| 13 | High quality sequence start: 14 | | | |
| 14 | High quality sequence stop: 653. | | | |
| 15 | Location/Qualifiers | | | |
| 16 | 1. 668 | | | |
| 17 | /organism="Gossypium arboreum" | | | |
| 18 | /strain="AKA" | | | |
| 19 | /cultivar="8400" | | | |
| 20 | /db_xref="taxon:29729" | | | |
| 21 | /clone="GA_Ed0066A09f" | | | |
| 22 | /cson.lib="Gossypium arboreum 7-10 dpa fiber library" | | | |
| 23 | /tissue_type="Fibers isolated from bolls harvested 7-10 | | | |
| 24 | dpa" | | | |
| 25 | /lab_host="E. coli" | | | |
| 26 | /note="Vector: pBR-CMV; site_1: EcoRI; site_2: XhoI" | | | |
| 27 | BASE COUNT | 197 a | 131 c | 143 g |
| 28 | ORIGIN | | | 197 t |
| 29 | Query Match | 75.6% | Score 422: | DB 14: Length 668: |
| 30 | Best Local Similarity | 84.8% | Pred: No. 2.6e-113: | |
| 31 | Matches 4/3; Conservative | 0; | Mismatches 85; | Indels 0; Gaps 0; |
| 32 | 1 | TGATGTGCACCCATTTATTTATGATAGCATGATATTCCTCAAAATTTGTGAACGACGTGC | 60 | |
| 33 | 55 | TGATTTTACACACTGTTTATGATACCGATATCTTCTCAAAATTTGTGAACGATGC | 114 | |
| 34 | 61 | GCCAAATTTGGAATACGTCCTATTGTACTGGAATTTATGCCATTATTAATTACAGG | 120 | |
| 35 | 115 | GCCAAATTTGGAATTTATTTGTCATGTCGCCGGTATTTATGCCATTATTAATACAGG | 174 | |
| 36 | 121 | GCATTTCCGCATGTCGCTTTTGGCAAAACAAAGATACGAGCTGATATGAGTGTCT | 180 | |
| 37 | 175 | GTTCTTGGCATGTCGCTTTTGGCAAAACAAAGATACGAGCTGATATGAGTGTCT | 234 | |
| 38 | 181 | TAGAGCCTTCAAGCAAAATGAGAGAGCTGTAAGCTTATGGAATTCACCTGGAGCTG | 240 | |
| 39 | 235 | TAGAGCCTTCAAGCAAAATGAGAGAGCTGTAAGCTTATGGAATTCACCTGGAGCTG | 294 | |
| 40 | 241 | AAATGTGCGCAAAATTTTATGCTATGATGATTAACATTCGATCTTTTACATCAATTA | 300 | |
| 41 | 295 | AAATGTGCGCAAAATTTTATGCTATGATGATTAACATTCGATCTTTTACATCAATTA | 354 | |
| 42 | 301 | TGGAGAAATCTGATCTGCAATCTATTAATACCTTGGGCTTAATGGAAGCTCAAGTT | 360 | |
| 43 | 355 | TGGAGAAATCTGATCTGCAATCTATTAATACCTTGGGCTTAATGGAAGCTCAAGTT | 414 | |
| 44 | 361 | CTAGAGCCTTACTTGGAGAGCGCGCAAGTCTTCCGTTAAAGAAAGATGTGCGTC | 420 | |
| 45 | 415 | CTAGAGCCTTACTTGGAGAGCGCGCGCAAGTCTTCCGTTAAAGAAAGATGTGCGTC | 474 | |
| 46 | 421 | CATCTTTTGGCAATATGACCAAAAGCTACATATCAAGAGCCATTTGATGGATCAAT | 480 | |
| 47 | 475 | CATCTTTTGGCTACACGCTCTTAAGGATATCAATTAACAGAGCAATTTGGTGGATCACT | 534 | |
| 48 | 481 | ACCCACATGGCGGTGGGTGATCTCTGTAATCCATCATATGTTGCATTAATGTATATC | 540 | |
| 49 | 535 | ACCCATTAATGGCGGTGGGTGATCTCTGTAATCCATCATATGTTGCATTAATGTATATC | 594 | |
| 50 | 541 | AGTTCATGGCGGCACGCTG | 558 | |

| Db | 595 | AGTCAATGAGCGCTCGAG | 612 |
|------------|---|---|----------------------------------|
| RESULT 9 | | | |
| LOCUS | BF269445 | 589 bp | linear |
| DEFINITION | GA_EB0004L07f Gossypium arboreum 7-10 dpa fiber library Gossypium | | |
| ACCESSION | BF269445 | | |
| VERSION | BF269445 | | |
| KEYWORDS | EST | | |
| SOURCE | Gossypium arboreum. | | |
| ORGANISM | Gossypium arboreum. | | |
| REFERENCE | Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotyledons II; Malvales; Malvaceae; Gossypium. | | |
| AUTHORS | Wang, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry | | |
| TITLE | 'D, Wood, T.C., Leslie, A. and Wilkins, T.A. | | |
| JOURNAL | An integrated analysis of the genetics, development, and evolution | | |
| COMMENT | Unpublished (2000) | | |
| FEATURES | Source | | |
| | 1..589 | | |
| | location/Qualifiers | | |
| | organism="Gossypium arboreum" | | |
| | strain="AKA" | | |
| | cultivar="8400" | | |
| | db_xref="taxon:29729" | | |
| | clone="GA_EB0004L07f" | | |
| | clone_lib="Gossypium arboreum 7-10 dpa fiber library" | | |
| | clone_type="fibers isolated from bolls harvested 7-10 dpa" | | |
| | note="host='E. coli'" | | |
| | vector="Vector: PBK-CMV; Site_1: EcoRI, Site_2: XhoI" | | |
| BASE COUNT | 170 a 112 c 125 g 180 t | 2 others | |
| ORIGIN | | | |
| | Query Match | 71.8%; | Score 400.4; DB 12; Length 589; |
| | Best Local Similarity | 85.3%; | Pred. No. 5.7e-107; |
| | Matches 446; Conservative | 0; | Mismatches 77; Indels 0; Gaps 0. |
| QY | 1 | TGATGTGACCCAAATTAATTTATGATACGGAATATATTCCTCAATATTTGGAACGACGTGC | 60 |
| DB | 66 | TGATTTGTACCAACGCTTTATATATACGATATATTTCTCAATTTGCAATGATATGTC | 125 |
| QY | 61 | GCCAAATTTGGAATACGCTGCTCTATATACGCTGGAATATATGCCATTAAATTTCAAGG | 120 |
| DB | 126 | GCCAAATTTGGAATTAATTTGCTCTATATGCTGCGGATATATGCCATTAAATTTCAAGG | 185 |
| QY | 121 | GCTTATGCCCATGACGTGGTGTGCAAAACAAAGATATACGACGTGACATTTAGCGTGTCT | 180 |
| DB | 186 | GTTCCTTTCGCAATAGATGGCTTTTGCAAAATCATAGATATACGACGTGATATACGTGCTC | 245 |
| QY | 181 | TGAGGCTATCAAGCAATGAACACCTGCTCAAGCTTATGAAATTCACGCTGCACTG | 240 |
| DB | 246 | TGAGGCTATCAAGCAACCAACCAACCTGCTCAAGCTTATGAAATTCACGCTGCACTG | 305 |
| QY | 241 | AAATGTGCAAAAGATTTACCTCATGGAATTTAAGCAATTCATCTTTATACATGAATA | 300 |
| DB | 306 | AAATGTGCGCAAGATTTTACCTCATGGAATTTAAGCAATTCATCTTTATACATGAATA | 365 |
| QY | 301 | TGAGGAATTCGATTTGCAATCTATATGACCTTGCGCTAATGGAAGTCCAAAGTTT | 360 |

| | | | |
|----|-----|--|-----|
| QY | 201 | GAAACACTGCAAGCGTTATNGCAATTCACCTGGGAACCTAATGTGCAAAAAAGATTTA | 260 |
| | | | |
| Db | 195 | GAGGAGCGCTGTGAAGAAGCATGTGCATCACCTGGTACTAGATGTGCCAAGAAAATTTTTG | 254 |
| QY | 261 | GCTCATTGGAATTAGACATTCGATCCTTTATACACTAAATATGAGCAAACTGCAATGGCA | 320 |
| | | | |
| Db | 255 | GCTAGTGGGATCAAGACTTTTGCACTTTTACACACTAAACATGTGAAAAAGACTGCTTAGCA | 314 |
| QY | 321 | ATACTAATGAACCTTGGCCCTAATGAAGAGCCCAAAGTCTCTAGTCCTTACCTTGGAGA | 380 |
| | | | |
| Db | 315 | ATTGTGATGAATCTTGGATTAATATAGAGAGCCCAAGCTTTCAAGAACATTACCTTGAGG | 374 |

| QY | 381 | CGCCCTGCAATGTTTTCGCGTAAAGAAGATCCGCTCCAACTTTTGGGCAAAACGA | 440 |
|----|-----|--|-----|
| QY | 375 | CCCTCAACTATGTTTTCCGCGTCAAGAGAGATGTTGCGCTATATTTCTGGGCAACAGA | 434 |
| Db | 441 | CCAAAAAGCTCAATCAATGAGGACATGAGTGGATCAATATACCACATGGGCGTTGGGGT | 500 |
| QY | 435 | CCAAAGATTACATCTCAAGAGCCACCTGGTTGGATGATGATCCACATGGGCGTGGGGT | 494 |
| Db | 501 | GATTCCTGTATTCATCATCATATGCTGTCATTTATCTGATTATCATGCTCATCGGACCACTG | 558 |
| QY | 495 | GATTCACAGAACCCATCATATATGAGGACCTTAATATACCAACAGATTTACACAGGACCGG | 552 |
| Db | 501 | | |

| RESULT 15 | LOCUS | DEFINITION |
|-----------|--|-------------------|
| BM526290 | 569 bp | linear |
| BM526290 | sa319604.y1 Gm-c1059 glycine max cdna clone | EST 19-FEB-2002 |
| | Gm-c1059-4400 5' similar to TR:Q9SU00 Q9SU00 | SOYBEAN CLONE ID: |
| | METHYLENETETRAHYDROFOLATE REDUCTASE ; | mRNA sequence. |

| | | |
|-----------|------------|-------------|
| ACCESSION | BM526290 | |
| VERSION | BM526290.1 | GI:18731158 |
| KEYWORDS | EST. | |
| SOURCE | soybean. | |

CONGRUENT TAXA
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS
1 (bases 1 to 569)
Shoemaker, R., Kelm, P., Voklin, L., Expelling, J., Corvelli, V., Khanna
, A., Bolla, B., Marry, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
, J., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
, R., Ritter, E., Kohn, S., Shiff, T., Jackson, I., Cardenas, M., Mcann
, R., Waterston, R., and Wilson, R.

| TITLE JOURNAL COMMENT | Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project |
|-----------------------------|---|
|-----------------------------|---|

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibdo
High quality sequence stop: 422.
Location/Qualifiers
1. .569

```

1..569
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl059-4400"
/clone_lib="Gm-cl059"
/tissue_type="Whole seedling, 2 week old, etiolated"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+, Site-1: EcoRI, Site-2:
XhoI; The cDNA library was constructed from mRNA isolated

```

from 2 week old etiolated whole seedlings of p14668916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI- XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."

| | Query Match | Similarity | Best Local | Matches | Conservative | Mismatches | Indels | Gaps |
|----|-------------|------------|--------------------|---------|--------------|------------|--------|------|
| QY | 1 | 95.3% | Score 364.6; | DB 13; | Length 569; | | | |
| Db | 190 | 97.4%; | Pred. No. 1.9e-96; | | | | | |
| QY | 61 | | | | | | | |
| Db | 230 | | | | | | | |
| QY | 121 | | | | | | | |
| Db | 310 | | | | | | | |
| QY | 181 | | | | | | | |
| Db | 370 | | | | | | | |
| QY | 241 | | | | | | | |
| Db | 430 | | | | | | | |
| QY | 301 | | | | | | | |
| Db | 490 | | | | | | | |
| QY | 361 | | | | | | | |
| Db | 550 | | | | | | | |

Search completed: February 14, 2003, 21:53:04
Job time : 2202 secs


```

; PRIOR APPLICATION NUMBER: US 08/738, 000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 660
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-931-795-2

```

```

Query Match      44.8%; Score 979; DB 9; Length 660;
Best Local Similarity 44.9%; Pred. No. 7.2e-85;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

```

```

QY 1 ARVITQLEFYDIDFLKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTRIPADIM 60
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 226 AFIITQLEFADTFPRFVKACDGMGICPIVPGIFPIQGYHSLRQLVKLSKLEVPQEK 285
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 AALEPINDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 119
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 286 DVIEPIDNDNAIRNGIELAVSLCOELASGLVGLHFYLLNKKEMATTEYLKRLGMWTE 345
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 SKVRSLLPWRPAPNVRVEEDVRPIFMANRPKSYISRTIGMDQYPHGRMGDSNPSYAL 179
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 346 DP-RRPLPMAISAPKREEDVRPIFMASSRPKSYITRQEWDEDPNGMSSSPAFGEL 404
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 180 SDVQ--FMRPRADKIVEMAVPLKSVEDIYERRLYCLGKLSN-----PMSELDG 230
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 405 KDYLFLKSKSKEELLMWGEELTSEASVEFVLLTSCDPNRNKHKTCTLPWMD-EP 463
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 231 LQPETKILNOLKIKNTKGLTINSOPAVNGEKSDSPVWGSGGYVYQKAYVEEFCCK 290
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 464 LAETSLKLELLRVNRQGLITINSQPNNGKPSDPIVWGSPGGYVFOKAYLEFFTSR 523
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 291 EKLDALVDKCD-RISLTYMAVNNKDSKSNVCGTDVNAVWGVFPKKEIIOPTIYDPS 349
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 524 ETAEALQVLAKYKELRVNHLVNVGENTNAPELQPMNAVWGIFFPGRREIIOPTVDPVS 583
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 350 FNVKDEAFETISRGMAVSLYPEDEASRKLYEEVSGSHFLVLYNDY-INGDLFAVPAD 407
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 584 FMEWKDEAFALMIERWKGKLYEEESPRTIYIHNDYFLVNLVNDPDLNCLMWVED 642
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 5

```

US-09-728-910-2
; Sequence 2, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Rima
; APPLICANT: Sekhon, Jaspreet
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETHETRAHYDROFOLATE
; TITLE OF INVENTION: REDUCTASE AND USES THEREOF
; FILE REFERENCE: 04844/006001
; CURRENT APPLICATION NUMBER: US/09/728, 910
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/258, 928
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 660
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-728-910-2

```

```

Query Match      44.8%; Score 979; DB 10; Length 660;
Best Local Similarity 44.9%; Pred. No. 7.2e-85;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

```

```

QY 1 ARVITQLEFYDIDFLKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTRIPADIM 60
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 226 AFIITQLEFADTFPRFVKACDGMGICPIVPGIFPIQGYHSLRQLVKLSKLEVPQEK 285
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 AALEPINDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 119
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 286 DVIEPIDNDNAIRNGIELAVSLCOELASGLVGLHFYLLNKKEMATTEYLKRLGMWTE 345
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 SKVRSLLPWRPAPNVRVEEDVRPIFMANRPKSYISRTIGMDQYPHGRMGDSNPSYAL 179
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 346 DP-RRPLPMAISAPKREEDVRPIFMASSRPKSYITRQEWDEDPNGMSSSPAFGEL 404
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 180 SDVQ--FMRPRADKIVEMAVPLKSVEDIYERRLYCLGKLSN-----PMSELDG 230
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 405 KDYLFLKSKSKEELLMWGEELTSEASVEFVLLTSCDPNRNKHKTCTLPWMD-EP 463
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 231 LQPETKILNOLKIKNTKGLTINSOPAVNGEKSDSPVWGSGGYVYQKAYVEEFCCK 290
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 464 LAETSLKLELLRVNRQGLITINSQPNNGKPSDPIVWGSPGGYVFOKAYLEFFTSR 523
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 291 EKLDALVDKCD-RISLTYMAVNNKDSKSNVCGTDVNAVWGVFPKKEIIOPTIYDPS 349
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 524 ETAEALQVLAKYKELRVNHLVNVGENTNAPELQPMNAVWGIFFPGRREIIOPTVDPVS 583
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 350 FNVKDEAFETISRGMAVSLYPEDEASRKLYEEVSGSHFLVLYNDY-INGDLFAVPAD 407
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 584 FMEWKDEAFALMIERWKGKLYEEESPRTIYIHNDYFLVNLVNDPDLNCLMWVED 642
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 6

```

US-10-113-852A-2
; Sequence 2, Application US/10113852A
; Publication No. US20020192784A1
; GENERAL INFORMATION:
; APPLICANT: Appling, Dean R.
; APPLICANT: Hanson, Andrew D.
; APPLICANT: Raymond, Rhonda R.
; APPLICANT: Roje, Sanja
; TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast S.
; FILE REFERENCE: 119927-1033
; CURRENT APPLICATION NUMBER: US/10/113, 852A
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,333
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 599
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
US-10-113-852A-2

```

```

Query Match      35.6%; Score 778; DB 9; Length 599;
Best Local Similarity 40.9%; Pred. No. 8.3e-66;
Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

```

```

QY 4 IVTQLEFYDIDFLKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTRIPADIM 63
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 185 IITQLEFYDIDFLKFNDCRQIGITCPIVPGIMPINNYKGFIRMTGCKTRIPADIM 244
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 64 EPIQDNEAKAYGILHGTGEMCKKILAHG-IKTLHLTYLNNKESALAILNNLGLIEE 121
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 245 DPLQDDELVDIGITNLYMCKCKLDSYVSHLTYLNNKESALAILNNLGLIEE 304
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 VSRS---LPWRPAPNVRVEEDVRPIFMANRPKSYISRTIGM--DOYPHGRMGDSNPS 175
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 305 FNAHPLAVLPPRRKSLNPKRNEVRPIFWRKRPYSYARTSQMVADEPPNRCFDSSPA 364
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 176 YGAL----SDYQPMRPRADKIVEMAVPLKSVEDIYERRLYCLGKLSNPMSELDO 231
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 365 FGDDLDLGGSD--ILKQSA--NCKLELWSTP--TSINDVAFVILNTINGKLCLEPSSDP-1 418
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 232 QPETKILNOLKIKNTKGLTINSOPAVNGEKSDSPVWGSGGYVYQKAYVEEFCCK 291
  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```
Db 419 NDEINPKAHLELNHNSITITNSQPVNGISNKGWPKDGYVQKQYLEFMKPKT 478
Qy 292 KLDLAIYKCKDRSTLYTMVANNKDGSKSN-VQOTVNAVTVGVPAPKEIIPTIVDVSR 350
Db 479 KLPKLTPLTKNNKEPLTFALDSQGDILSNHPNSKSNMVTWGIPEGREIIQPTIVEKISF 538
Qy 351 NWKDEAFETWSRGMSLYPEDEA-SRKLVEVSGSHPLVSLVDNDYIND 400
Db 539 LAKKEEYHILNEMKLNMMNKYDKPHGACFIQSLIDYCLVMIYNDYISPD 589
```

RESULT 7

```
US-09-734-017A-62
; Sequence 62, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lersch, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwend, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 62
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-017A-62
```

```
Query Match 33.9%; Score 742; DB 10; Length 191;
Best Local Similarity 79.9%; Pred. No. 4,1e-63;
Matches 135; Conservative 14; Mismatches 20; Indels 0; Gaps 0;
```

```
Qy 3 VIVTQLEFYDDIFLKFVNDRCRQIGTCPIVPGIMPINNYKGFIRMTGFCFKIIPADIMAA 62
Db 3 VIVTQLEFYDDIFLKFVNDRCRQIGTCPIVPGIMPINNYKGFIRMTGFCFKIIPADIMAA 62
Qy 63 LEPTKDEEAVKAVYGIHGTGEMCKKTLAHGKTLHLTYLTLNNEKSAIILMMGLIENSKY 122
Db 63 LEPTKDEEAVKAVYGIHGTGEMCKKTLAHGKTLHLTYLTLNNEKSAIILMMGLIENSKY 122
Qy 123 SRSILPMRRPANNVFRKEDVAPLTFWANRPKSYISRTIGMDQYPHGRWDS 171
Db 123 SRSILPMRRPANNVFRKEDVAPLTFWANRPKSYISRTIGMDQYPHGRWDS 171
```

RESULT 8

```
US-09-734-017A-64
; Sequence 64, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lersch, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
```

```
; APPLICANT: Freund, Annette
; APPLICANT: Duwend, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 64
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-017A-64
```

```
Query Match 9.1%; Score 199; DB 10; Length 136;
Best Local Similarity 90.2%; Pred. No. 1.2e-11;
Matches 37; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 3 VIVTQLEFYDDIFLKFVNDRCRQIGTCPIVPGIMPINNYK 43
Db 87 VIVTQLEFYDDIFLKFVNDRCRQIGTCPIVPGIMPINNYK 127
```

RESULT 9

```
US-09-738-626-5879
; Sequence 5879, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MARGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5879
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5879
```

```
Query Match 7.2%; Score 158; DB 9; Length 326;
Best Local Similarity 27.7%; Pred. No. 3.3e-07;
Matches 36; Conservative 28; Mismatches 46; Indels 20; Gaps 4;
```

```
Qy 1 ARVIVTQLEFYDDIFLKF-----VNDRCRQIGTCPIVPGIMPINNYKGFIRMTGFC 51
Db 189 AYSISIQMFVEDVDYRLRLDRVLAADPIHGAK-----PIIPGIMPITSRSVRQVELS 242
Qy 52 KTKIPADIMAA-EPKONEA-----VKAYGIHGTGEMCKKTLAHGKTLHLTYLTLNMS 106
Db 243 GAQIPSOLESISIVRAANGNEANKDEIRKVGIEYSTNAERLIAEGADLHFTLNFTRA 302
```



```
QY      107 ALAILNLGL 116
          :| | | :
Db      303 TQEVLYNLGM 312
```

```

RESULT 10
US-09-919-935-2
/ Sequence 2, Application US/09919935
/ Patent No. US20020049305A1
GENERAL INFORMATION:
APPLICANT: BATHE, BRIGITTE
APPLICANT: MCKELL, BETTINA
APPLICANT: PFEFFERLE, WALTER
APPLICANT: HOTHMACHER, KLAUS
APPLICANT: BINDER, MICHAEL
APPLICANT: GREISSINGER, DIETER
APPLICANT: THIERBACH, GEORG
TITLE OR INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE METE GENE
FILE REFERENCE: 211/12850X
CURRENT APPLICATION NUMBER: US/09/919,935
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: DE 10053942.4
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: DE 10199666.0
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/294,279
PRIOR FILING DATE: 2001-05-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 349
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-919-935-2

```

| | | | | |
|-----------------------|--------|--------------------|----------------|-------------|
| Query Match | 7.2%; | Score 157; | DB 10; | Length 349; |
| Best Local Similarity | 27.7%; | Pred. No. 4.6e-07; | | |
| Matches | 36; | Conservative 27; | Mismatches 47; | Indels 20; |
| | | | | Gaps 4; |

[illegible]

1 RESULT 11
 2 US-09-943-702-5
 3 Sequence 5, Application US/09943702
 4 Publication No. US20030022322A1
 5
 6 GENERAL INFORMATION:
 7
 8 APPLICANT: Rostock, Bradley S.
 9 Rostock Jr., Paul R.
 10 TITLE OF INVENTION: SAM Operon
 11 NUMBER OF SEQUENCES: 6
 12 CORRESPONDENCE ADDRESS:
 13 ADDRESS: Eli Lilly and Company
 14 STREET: Lilly Corporate Center
 15 CITY: Indianapolis
 16 STATE: Indiana
 17 COUNTRY: U.S.
 18 ZIP: 46285
 19
 20 COMPUTER READABLE FORM:
 21 MEDIUM TYPE: Floppy disk
 22 COMPUTER: IBM PC compatible
 23 OPERATING SYSTEM: PC-DOS/MS-DOS

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/943,702
:   FILING DATE: 31-Aug-2001
:   CLASSIFICATION: <Unknown>
:
: PRIORITY APPLICATION DATA:
:   APPLICATION NUMBER: 08/955,957
:   FILING DATE: 1997-10-22
:
: ATTORNEY/AGENT INFORMATION:
:   NAME: Webster, Thomas D.
:   REGISTRATION NUMBER: 39,872
:   REFERENCE/DOCKET NUMBER: P-10162
:
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 317-276-3334
:
: INFORMATION FOR SEQ ID NO: 5:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 305 amino acids
:     TYPE: amino acid
:     TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 5:
: US-09-943-702-5

```

| | | | | |
|--------------------------|--------|--------------------|-----------|-------------|
| Query Match | 6.38; | Score 138; | DB 9; | Length 305; |
| Best Local Similarity | 27.38; | Pred. No. 2.4e-05; | | |
| Matches 30; Conservative | 18; | Mismatches 62; | Indels 0; | Gaps 0; |

```

OY      5 YQLFEDYDDIDFLKYNDCROJGTCIPVYPGPIINNKGFIRMGCEFKTIPADIMALE 64
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      179 ITOMFREVEDYIRLDRVAAAGCCPTVPEIMPATDVKQIRFELSHAFPEBLARLE 238
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      65 PIKDNEAVKAVGIIHLGTEMCKIIANGIKTLHLXTLMEKSAIILML 114
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      239 AARGNPAGCHRGIVYATAMAGRLLEAEPAGIIHTITNRSTATLEIHRNI 288
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12
 US-10-153-273-12
 ; Sequence 12, Application US/10153273
 ; Patent No. US20020169305A1
 ;
 GENERAL INFORMATION:
 ;
 APPLICANT: Sim, Kim L.
 ;
 Chinitz, Chetan
 ; Miller, Louis H.
 ; Peterson, David S.
 ; Su, Xin-zhan
 ; Wellem, Thomas E.
 ;
 TITLE OF INVENTION: AND PLASMIDIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 ;
 NUMBER OF SEQUENCES: 37
 ;
 CORRESPONDENCE ADDRESS:
 ;
 ADDRESSEE: Knobbe Martens Olson & Bear
 ;
 STREET: 620 Newport Center Drive 16th Floor
 ;
 CITY: Newport Beach
 ;
 STATE: California
 ;
 COUNTRY: US
 ;
 ZIP: 92660
 ;
 COMPUTER READABLE FORM:
 ;
 MEDIUM TYPE: Floppy disk
 ;
 COMPUTER: IBM PC compatible
 ;
 OPERATING SYSTEM: PC-DOS/MS-DOS
 ;
 SOFTWARE: Patentin Release #1.0, Version #1.25
 ;
 CURRENT APPLICATION DATA:
 ;
 APPLICATION NUMBER: US/10/153,273
 ;
 FILING DATE: 21-May-2002
 ;
 CLASSIFICATION: <Unknown>
 ;
 PRIOR APPLICATION DATA:
 ;
 APPLICATION NUMBER: US/09/210,288
 ;
 FILING DATE: <Unknown>
 ;
 ATTORNEY/AGENT INFORMATION:
 ;
 NAME: Fuller, Michael
 ;
 REGISTRATION NUMBER: 36,516
 ;
 REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
 ;

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-153-273-12

Query Match 4.5%; Score 97.5; DB 9; Length 2710;
Best Local Similarity 19.8%; Pred. No. 4.3;
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMAAAL-EPKIDNEAVKAGIHGTEM---CKILAHGKILHLTYLTMKESALAIL 111
DB 182 PSQCTWLARSFADIDIVRGHDLVGNPQETKROOLENNILKITL----- 226
QY 112 MNLGLIEESKVSRLPWRPRAVFRVKEDEVRIFW-ANRPKSYISRTIGMDQYPHGMGD 170
DB 227 --FGKIYEKINGAARVYGDNDPEEFKRED---WMTANR-----ETV-WKAITCNMAGN 273
QY 171 S-----CN-----PSYCALSDY--QFMRPRARDKKLVEEAVL----- 202
DB 274 TYFHATNCRGERTKGYCRCDNDQVPTV--FDYVPOYLK-----WFEWAEDFCRRKN 323
QY 203 KSVEDIYERFR-----LYC-----LGLRLR-----SNPW----- 225
DB 324 KIKIDVARNCRGKDEKEDRYGSRNGYDCEKTKRAIGLRGKQGISCLYACNRYVDWIN 383
QY 226 SELDGLQPEIKINLEOLEKINTKGLTITSQPAVGEKSD---SPTVGMGSPGGYVYOKA 282
DB 384 NQKQFQKQKKYDDEEKITE-----NGASGSRQKRDAGTTTNYDG-----YEKK 431
QY 283 Y-----VEFCSREKLDALVDCCKDRT--SLTYMAVNNKDSWKSNAVQOTDYNVAV 329
DB 432 FYDELANKSEYRTVDKFLKLSNNEICTKVKDEGGTIDFKNNSD---SISGASGATNVE 487
QY 330 TWGTFPAKKEITQP-----TIYDPSFVWAKDEAFETLSKGMASLY---PEDASRKLV 379
DB 488 SQGFFYSKSYCPCPYCGVAKVNNNGSSNMEWEEKNNGRCKSG--KLYPEPKDKEGTTITI 545
QY 380 EEVGSHPFLVSLVDN---DYINGD 400
DB 546 LKSGKHGHDIEKLNKFCDEKND 569

RESULT 13
US-09-911-888-23
Sequence 23, Application us/09911888
Patent No. US20020119509A1
GENERAL INFORMATION:
APPLICANT: Kaitlin, Yigal
APPLICANT: Gavrias, Victoria
TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
FILE REFERENCE: 06286-062002
CURRENT APPLICATION NUMBER: US/09/911,888
PRIOR FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/965,762
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 347
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-911-888-23

Query Match 4.2%; Score 91; DB 10; Length 347;
Best Local Similarity 21.3%; Pred. No. 0.87;
Matches 78; Conservative 57; Mismatches 118; Indels 114; Gaps 22;

QY 2 RVIYTOLEFYDTDLFLKVNDCRQIGITCPHYQIMINNYKQFIMTQPKTKPADIMA 61
DB 22 RLITE--DTNRLITVYTSR---TLPRVQEV--INQIKFYKSRVVE-DLEIDPDY 71
QY 62 ALEPKIDNEAVKAGIHGTEMCKKILAHGKILHLTYLTMKESALAIL-MN----- 113
DB 72 LLVDFTMNVSILNAY-----YQINKRYAIRNLFVMAAQGFED 109
QY 114 ---LGLIEESKVSRLPWRPRAVFRV-----KEVRIFFAN--RKSYISRTIGMDQ 162
DB 110 GIDWIGAVKEVETNPLEAVTNPYYKIQLVGYSKSDMDGLIFQANVGFYPFYSKITL--PQ 167
QY 163 YPHGR---WQDS--CNPSYCALSDYQFMRPRAR--DKLVEEAVPLKSVEDI----- 208
DB 168 LTRGKATYVWISSIMSDPKTSLNDIELLKTNSYEGSKRLVD--LHLATYKDLKTKGI 225
QY 209 -----YERFRLYC---LGLRSNPSELDGLOPE-----TKIT 238
DB 226 NQYVVOPIFTSHSFSSEYLNFFTFGMLCLFYLARLIGSPMHNIDGYKAANAPVYVTRIA 285
QY 239 NEOLEKINTK-GFLTI--NSQPAVNGEKSDSPVGMGSPGGYVYOKAYVEFCSREKLDAL 296
DB 286 NPNFEKQDVKGSATSRDGMPIKTQCID-----PTGMSDVFRATYO---KKKLE-W 332
QY 297 VDKCKDR 303
DB 333 DEKLKQ 339

RESULT 14
US-10-003-392-2
Sequence 2, Application us/10003392
Patent No. US20020170086A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen M.
APPLICANT: Caiml, Perry G.
APPLICANT: Stoop, Johan M.
TITLE OF INVENTION: Fructan Biosynthetic Enzymes
FILE REFERENCE: BR1463 US NA
CURRENT APPLICATION NUMBER: US/10/003,392
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,273
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: 60/269,543
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 608
TYPE: PRT
ORGANISM: Dimorphotheca sinuata
US-10-003-392-2

Query Match 4.2%; Score 91; DB 9; Length 608;
Best Local Similarity 18.9%; Pred. No. 2;
Matches 86; Conservative 69; Mismatches 150; Indels 150; Gaps 25;

QY 60 MALEPIKDNEAVKAGIHGTEMCKKILAH-----GIKTLHLTYLTMKESA 107
DB 1 MTTTKFESDLEDA-----PLNHTREPPPPPPPTAGRKLLIYVSVITLL 46
QY 108 LAILMNLGLIEESKVSRS-----LPWRPRAVFRVKEDEVRIFWANRPKSYIS 155
DB 47 ILLIYSVLFLNQONSSHSTNSKSISSQDRLIWEKTSFHFQPAKNF--IYDENGFLH-- 102
QY 156 RTIGW---DQY-PHGR-WQDSQCNPSYCALSDYQFMRPRARDKKLVEEAVV----- 201
DB 103 --MGWYHLFYQINPYGPWVG--NMSWG-----HSVSKDMINNEDELVALVPTFW 147

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 12, 2003, 17:03:12 : Search time 75 Seconds

(without alignments)
2770.721 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187
Sequence: 1 ARVYTGLEFYDTDFLKFN.....VSLVDNDYINGDLFAVADF 408

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 424239 segs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+P2n.model -DEV=Xlh
-O=/cgn2_1/USPTO.spool/US09720451/runat_05022003.074810_2410/app_query.fasta.1.583
-DB=Published_Applications_NA -OPM=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09720451.@cgn2.1.1.26.@runat_05022003.074810_2410
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEROUT=120 -WARN_TIMEROUT=30 -THRAPDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCU_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCUTS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | Score | Match | Length | ID | Description |
|--------|-------|-------|-------|--------|------------------|-------------------|
| No. | | | | | | |
| 1 | 1777 | 81.3 | 1994 | 9 | US-10-113-852A-3 | Sequence 3, Appl1 |
| 2 | 979 | 44.8 | 2219 | 9 | US-09-931-795-3 | Sequence 3, Appl1 |
| 3 | 979 | 44.8 | 2219 | 10 | US-09-728-910-3 | Sequence 3, Appl1 |
| 4 | 979 | 44.8 | 2220 | 9 | US-09-931-795-1 | Sequence 1, Appl1 |

| | | | | | | |
|----|------|------|--------|----|---------------------|--------------------|
| 5 | 979 | 44.8 | 2220 | 10 | US-09-728-910-1 | Sequence 1, Appl1 |
| 6 | 950 | 43.4 | 832 | 10 | US-09-770-445-690 | Sequence 690, App |
| 7 | 778 | 35.6 | 3082 | 9 | US-10-113-852A-1 | Sequence 1, Appl1 |
| 8 | 775 | 35.4 | 574 | 10 | US-09-734-017A-61 | Sequence 61, Appl1 |
| 9 | 398 | 18.2 | 313 | 10 | US-09-878-574-3765 | Sequence 3765, Ap |
| 10 | 390 | 17.8 | 264 | 10 | US-09-923-876-5596 | Sequence 5596, Ap |
| 11 | 389 | 17.8 | 265 | 10 | US-09-923-876-476 | Sequence 476, App |
| 12 | 284 | 13.0 | 181 | 10 | US-09-923-876-5753 | Sequence 5753, Ap |
| 13 | 274 | 12.5 | 216 | 10 | US-09-923-876-1503 | Sequence 1503, Ap |
| 14 | 255 | 11.7 | 159 | 10 | US-09-294-0938-4653 | Sequence 4653, Ap |
| 15 | 213 | 9.7 | 409 | 10 | US-09-734-017A-63 | Sequence 63, Appl1 |
| 16 | 168 | 7.7 | 640681 | 9 | US-09-780-888-1 | Sequence 1, Appl1 |
| 17 | 158 | 7.2 | 378 | 9 | US-09-738-626-2379 | Sequence 2379, Ap |
| 18 | 157 | 7.2 | 1551 | 10 | US-09-919-835-1 | Sequence 1, Appl1 |
| 19 | 138 | 6.3 | 4848 | 9 | US-09-943-702-1 | Sequence 1, Appl1 |
| 20 | 138 | 6.3 | 4848 | 9 | US-09-943-702-4 | Sequence 4, Appl1 |
| 21 | 138 | 6.3 | 4848 | 9 | US-09-943-702-6 | Sequence 6, Appl1 |
| 22 | 100 | 4.6 | 38584 | 9 | US-10-114-170-50 | Sequence 50, Appl1 |
| 23 | 97.5 | 4.5 | 8220 | 9 | US-10-153-273-11 | Sequence 11, Appl1 |
| 24 | 93.5 | 4.3 | 1900 | 10 | US-09-911-888-22 | Sequence 22, Appl1 |
| 25 | 93.5 | 4.3 | 1900 | 10 | US-09-911-888-24 | Sequence 24, Appl1 |
| 26 | 91 | 4.2 | 2080 | 9 | US-10-003-382-1 | Sequence 1, Appl1 |
| 27 | 89.5 | 4.1 | 7053 | 10 | US-09-070-927A-78 | Sequence 1, Appl1 |
| 28 | 88 | 4.0 | 1182 | 9 | US-10-071-338-6 | Sequence 6, Appl1 |
| 29 | 88 | 4.0 | 3364 | 10 | US-09-925-300-665 | Sequence 665, App |
| 30 | 88 | 4.0 | 7193 | 9 | US-10-071-338-1 | Sequence 1, Appl1 |
| 31 | 87.5 | 4.0 | 1610 | 10 | US-09-954-456-92 | Sequence 92, Appl1 |
| 32 | 87.5 | 4.0 | 1610 | 10 | US-09-967-768A-219 | Sequence 219, App |
| 33 | 87.5 | 4.0 | 2393 | 10 | US-09-964-824A-254 | Sequence 254, App |
| 34 | 87.5 | 4.0 | 2433 | 9 | US-09-738-626-2317 | Sequence 2317, Ap |
| 35 | 87.5 | 4.0 | 4765 | 12 | US-10-090-624-5 | Sequence 51, Appl1 |
| 36 | 87 | 4.0 | 2956 | 10 | US-09-895-382-31 | Sequence 31, Appl1 |
| 37 | 85.5 | 3.9 | 1329 | 9 | US-09-738-626-2703 | Sequence 2703, Ap |
| 38 | 85.5 | 3.9 | 1674 | 9 | US-09-738-626-2702 | Sequence 2702, Ap |
| 39 | 85.5 | 3.9 | 6449 | 9 | US-09-938-882A-69 | Sequence 69, Appl1 |
| 40 | 85 | 3.9 | 963 | 10 | US-09-070-927A-547 | Sequence 547, App |
| 41 | 85 | 3.9 | 963 | 10 | US-09-927-091-3 | Sequence 3, Appl1 |
| 42 | 85 | 3.9 | 6544 | 9 | US-10-004-381-40 | Sequence 40, Appl1 |
| 43 | 85 | 3.9 | 6544 | 9 | US-10-004-381-40 | Sequence 37, Appl1 |
| 44 | 84.5 | 3.9 | 1846 | 10 | US-09-823-356-24 | Sequence 24, Appl1 |
| 45 | 84.5 | 3.9 | 2070 | 12 | US-10-044-090-377 | Sequence 377, App |

ALIGNMENTS

RESULT 1
US-10-113-852A-3

Sequence 3, Application US/10113852A
Publication No. US20020192784A1

GENERAL INFORMATION:

APPLICANT: Applling, Dean R.
APPLICANT: Hanson, Andrew D.
APPLICANT: Raymond, Rhonda R.

APPLICANT: Roje, Sanja
TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast S.

FILE REFERENCE: 119927-1033

CURRENT APPLICATION NUMBER: US/10/113, 852A

PRIOR FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: 60/280,333

NUMBER OF SEQ ID NOS: 13

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 1994

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: CDS

LOCATION: (63)..(1841)

OTHER INFORMATION: DNA construct

US-10-113-852A-3

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2.16e-197 | Length: | 1994 |
| Score: | 1777.00 | Matches: | 319 |
| Percent Similarity: | 89.41% | Conservative: | 44 |
| Best local Similarity: | 78.57% | Mismatches: | 43 |
| Query Match: | 81.25% | Indels: | 0 |
| GB: | 9 | Gaps: | 0 |

US-09-720-451-6 (1-408) x US-10-113-852A-3 (1-1994)

| | | | |
|----|------|--|-------|
| Oy | 1 | AlaAcgValIleValThrGlnLeuPhePheYAspPrrAspIlePheLeuPheValAsn | 20 |
| | | | |
| Db | 621 | GCGAGTTTCAATTGTCACACGCTTTCTGTACTGATATATATTCCTCAAGTTTGGAAT | 680 |
| Oy | 21 | AspCysAlaGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn | 40 |
| | | | |
| Db | 681 | GATTTCGGCAAAATCGGATTAATATTCGCCATATGTTCCGGAAATATAGCCATTAATCAAC | 740 |
| Oy | 41 | TyrIlycGlyPheIleArgMetThrGlyPheCysLysThrIlyIleProIleAspIleMet | 60 |
| | | | |
| Db | 741 | TACAAAGGGTCTTCTCCGTAATAGCGTGTTCTGTAAAGACCAAGATACCCGCTGACCTACT | 800 |
| Oy | 61 | AlaAlaIleuGlnProIleLysAspAsnGluGluAlaValLysAlaIlyrGlyIleHisLeu | 80 |
| | | | |
| Db | 801 | GCGCCTTAGACCCATTAAGATATGACGAGGCGTGTAAAGCCTATGGAATTCACCTT | 860 |
| Oy | 81 | GlyThrGlnMetCysLysIlyIleLeuAlaHisGlyIleLysThrIleuHisLeuThr | 100 |
| | | | |
| Db | 861 | GCAACGAAATGTGCAGAAAGATATTTGGCCACGGAATACCTCCCTCACTCTCAACA | 920 |
| Oy | 101 | LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlnSer | 120 |
| | | | |
| Db | 921 | TTTCAAGCTGGACAAATACGATTATGGATATTAAGAACCTGGCTCGATGTATGATCA | 980 |
| Oy | 121 | LysValSerArgSerLeuProIleProIleArgArgProAlaAsnValPheArgValLysGluAsp | 140 |
| | | | |
| Db | 981 | AAATATTTCTGTTCTTCACCTTGAGACCCCTCGCAATGTTTTCCGTAATAGGAAGAT | 1040 |
| Oy | 141 | ValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyThr | 160 |
| | | | |
| Db | 1041 | GTTCCGCCAATTTCTGGGCAACCGTCCAAAGACCTCAATRTGTAGAACAAAGGCGTGG | 11000 |
| Oy | 161 | AspGlnTyrProHisGlyAlaGlyTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer | 180 |
| | | | |
| Db | 1101 | AATGACTTCCCACTGACGACGTGGGGGATGTACACAGATGAGATCACTACACTCTTGG | 1160 |
| Oy | 181 | AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTrpAlaVal | 200 |
| | | | |
| Db | 1161 | GATTATCACTTTGGCGCCGCAAAAGACGCTGACCAAGAGCTTCAGCAAGATGGGCGTC | 1220 |
| Oy | 201 | ProLeuLysSerValGluAspIleTyrGluArgPheArgGlyLeuTyrCysLysGlyLysLeu | 220 |
| | | | |
| Db | 1221 | CCACTGAAAAGCATGTAGAGATGTCACAGAAATTCAAAGAGCGTCTCATGTGAAACTTA | 1280 |
| Oy | 221 | ArgSerAspProTrpSerGlnLeuAspGlyLeuGlnInProGluThrIlyIleHisGln | 240 |
| | | | |
| Db | 1281 | AAAGACAGCCCATGTGTCGATTATGATGACATCCAGCCAGAGACAAAGATCATTAATGAG | 1340 |
| Oy | 241 | GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn | 260 |
| | | | |
| Db | 1341 | CAACTCGGAAAAATCAACTCCACAGCGTTTCCGACCATTAATACCAACCATAGTCAT | 1400 |
| Oy | 261 | GlyGlnLysSerAspSerProThrValAlaGlyTrpGlyGlyProGlyGlyTyrValTyrGln | 280 |
| | | | |
| Db | 1401 | GCGAGCCAATTCGATTTCTCCAGCATATGAGATGGGGGGTCTCGTGGTACGTCACAG | 1460 |
| Oy | 281 | LysAlaIleTyrValGlnPhePheCysSerLysGlyIleAsnAlaIleuValAspLysCys | 300 |
| | | | |
| Db | 1461 | AAAGCTTATCTGGAATTTCTGTGTCAAAAGATTAAGTTAGACACACTTGGGAGAAATCC | 1520 |
| Oy | 301 | LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTyrLysSerAsn | 320 |
| | | | |
| Db | 1521 | AAAGCTTTTCCCTTCATATCACTCAATAGCGCGGAAACAATAGAGAATTTGGATATCAAC | 1580 |

| | | | |
|----|------|---|------|
| QY | 321 | ValISGIshThsApvAlAsnAlaVtThTPrGIVaLpHePrAlaYssGluIle | 340 |
| | | | |
| Db | 1581 | ACCGGAGATCCGATGTAAAGCAGTACTCTGGGAGGTCCACGTAAGAGGTATT | 1640 |
| QY | 341 | GIuPrOthLlIeVaLAsPrOvaLSePrheAsVaLTrpLYsAspGluAlaPheGluIle | 360 |
| | | | |
| Db | 1641 | CAACCAACATGTGTGATACCGCAGCTTTTAAAGTGTGGAAAGACAGCGTTTGAATT | 1700 |
| QY | 361 | TrpSerArgLYTPALaSerLeuTYrPrOGluAspGluAlaSerArgLYLeuValGlu | 380 |
| | | | |
| Db | 1701 | TGGCGAAGAGCTGGCTACCTTTTACCACGAAGATGACCTCTTCAAAAGTGTCTCGAG | 1760 |
| QY | 381 | GIuValGlyLYSerHisPheLeuValSerLeuValAspAsnAspTYrLlIeAsnGlyAsp | 400 |
| | | | |
| Db | 1761 | GAGGTAAAGAACAGCTACTATTATTGGTAAAGCTAGTGCAACAACATTACATCAATGGTGTAT | 1820 |
| QY | 401 | LeuPheAlaValPheAla | 406 |
| | | | |
| Db | 1821 | ATCTTCTGTCTGTTCCT | 1838 |

RESULT 2

US-09-931-795-3
; Sequence 3, Application US/09931795

Publication No. US20020198211A1
; GENERAL INFORMATION:

APPLICANT: ROZEN, Rima
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE

FILE REFERENCE: 04844/005003

```

; CURRENT APPLICATION NUMBER: US/09/931,795
; CURRENT FILING DATE: 2001-08-16

```

;; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12

;; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01

;; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12

;; PRIOR APPLICATION NUMBER: PCT/CA95/0031
; PRIOR FILING DATE: 1995-05-25

;; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4

```

```

; SEQ ID NO 3
; LENGTH: 2219

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

```
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: (13)...(1983)
US-09-931-795-3

Alignment Scores:

| | | |
|------------------|----------|-----|
| Pred. NO.: | 3.09E+04 | Len |
| Score: | 979.00 | Mat |
| Percent Correct: | 63.73% | Con |

| | | |
|------------------------|--------|------|
| Percent Similarity: | 83.728 | com |
| Best Local Similarity: | 44.878 | Miss |
| Customer Match: | 44.768 | Trac |

| | | |
|--------------|--------|-----|
| query match: | 44.708 | Inc |
| DB: | 9 | Gap |

US-09-720-451-6 (1-408) x US-09-931-795-3

QY I AlaArgvalIlevalThrgInLeuPheTyrAs

Db 676 GCCGATTTCATCATCAGCAGCTTTCCTTGA

QY 21 AspCysArgGlnIleGlyIleThrCysProI

Db 736 GCATGCACCGACATGGGCATCACTTGCCCAT

QY 41 TYRGLYSGLYPHEILEARGMETTHRGLYPHECY

Db 796 TACCACCTCCCTCGGCAGCTGTGAAGCTGTC


```
Db 1213 AAGGACTACTACCTTCTTACCTGAAGAGCAAGTCCGCCAAGAGAGAGCTGCTGAAGATG 1272
Qy 198 TTPAlaValProLeuYsSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGCTGACACAGTGAAGCAAGTCTTGAAGCTTTGCTTACCTCTCG 1332
Qy 218 GLYIysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCGGAATGCTCAAAAGTACTGCTGCTGCTGCTGCAACAT---GAGCCC 1389
Qy 231 LeuGlnProGluThrIleLeuLeuGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGln 250
Db 1390 CTGGGGCTGAGACACAGCTGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
Qy 231 LeuThrIleAsnSerGlnProAlaValAsnGlnIleuYsSerAspSerProThrValGly 270
Db 1450 CTCACATCAACATCAACAGCCCAACATCAACAGGAGAGAGAGAGAGAGAGAGAGAG 1509
Qy 271 TTPelGlyProGlyIleTyrValTyrGlnIleuYsAlaTyrValGluPhePheCysSerIys 290
Db 1510 TGGGGCCCAAGCGGGGCTATGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
Qy 291 GluIleuYsSerAspIleuValAspIleuYsCysIysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
Qy 310 AlaValAsnIysAspGlySerTyrIleuValGlnIleuThrAspValAsnAlaVal 329
Db 1630 CTTCATCATGTGAAGGTGAAGAAATCAACATCAACATGCCCCCTGACAGCCGATGCTGTC 1689
Qy 330 ThrTTPelGlyValPheProAlaIysGlnIleuLeuGlnIleuProThrIleValAspProValSer 349
Db 1690 ACTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1749
Qy 350 PheAsnValIleTyrIysAspGluAlaPheGlnIleuTyrSerArgIleTyrAlaSerIleuTyr 369
Db 1750 TTTCATGTTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1809
Qy 370 ProGluAspGluAlaSerArgIleuValGlnIleuValGlnIleuValGlnIleuVal 389
Db 1810 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1869
Qy 390 SerLeuValAspAsnAspTyr---IleAsnGlnIysPheLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGGGAGACATGACTTCCCACTGGACAACTGCTGCGAGAGTGGTGAAGAC 1926
```

RESULT 4

```
us-09-931-795-1
; Sequence 1, Application US/09931795
; Publication No. us20020198211A1
; GENERAL INFORMATION:
; APPLICANT: ROZEN, Rima
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETHETRAHYDROFOLATE
; FILE REFERENCE: 04844/005003
; CURRENT APPLICATION NUMBER: US/09/931,795
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 09/592,595
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/738,000
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: PCT/CA95/00314
; PRIOR FILING DATE: 1995-05-25
; PRIOR APPLICATION NUMBER: GB 9410620.0
; PRIOR FILING DATE: 1994-05-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
us-09-931-795-1
```

Alignment Scores:

| Pred. No.: | 3 09e-104 | Length: | 2220 |
|------------------------|-----------|---------------|------|
| Score: | 979.00 | Matches: | 188 |
| Percent Similarity: | 63.72% | Conservative: | 79 |
| Best Local Similarity: | 44.87% | Mismatches: | 138 |
| Query Match: | 44./68 | Indels: | 14 |
| Ds: | 9 | Gaps: | 7 |

US-09-720-451-6 (1-408) x US-09-931-795-1 (1-2220)

```
Qy 1 AlaValAlaValIleuThrGlnLeuPheTyrAspThrIlePheLeuYsPheValAsn 20
Db 676 GCCGATTCATCATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 735
Qy 21 AspCysArgGlnIleGlyIlePheCysProIleValProGlyIleMetProIleAsnAsn 40
Db 736 GCATGACACCGACATGGGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 795
Qy 41 TyrIysGlyPheIleArgMetThrGlyPheCysIysThrIysIleProAlaAspIleMet 60
Db 796 TACCACTCCCTTGGCAGCTTGTGAAGCTGTCCAAAGCTGGAGGTGCCACAGAGATCAAG 855
Qy 61 AlaAlaLeuGluProIleIysAspAsnGlnIleuAlaValIysAlaTyrGlyIleIleIleu 80
Db 856 GACGTGATGAGCAATCAAAAGACATGCTCATCCGCAACATCAATGATGATGATGATG 915
Qy 81 GlyThrIleuTyrCysIysIleLeuAlaIleGly---IleYsThrIleuIleuTyr 99
Db 916 GCCGTGACCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 975
Qy 100 ThrIleuAsnMetGluYsSerAlaLeuAlaIleuMetAsnLeuGlnIleuGlnIleu 119
Db 976 ACCCTCAACCGCAGATGCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1035
Qy 120 SerIysValSerArgSerLeuProTyrArgProAlaAsnValPheArgValIysGln 139
Db 1036 GACCCC---AGGCTGCCCTTACCTGAGGCTCTGAGTCCACCCCAAGAGAGAGAGAA 1092
Qy 140 AspValArgProIlePheThrPheAlaAsnArgProIysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTAGTCCATCTTCTTGGGCTCCAGACCAAGATTAATCTTACCGTACCCAGAGAG 1152
Qy 160 TTPAspIleTyrProIleGlnArgTTPelGlyAspSerCysAspProSerGlyAlaLeu 179
Db 1153 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1212
Qy 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspIysIysLeuValGlnIle 197
Db 1213 AAGGACTACTACCTTCTTACCTGAAGAGCAAGTCCGCCAAGAGAGAGAGAGAGAG 1272
Qy 198 TTPAlaValProLeuYsSerValGluAspIleTyrGlnArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1332
Qy 218 GLYIysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCGGAATGCTCAAAAGTACTGCTGCTGCTGCTGCAACAT---GAGCCC 1389
Qy 231 LeuGlnProGluThrIleLeuLeuGlnIleuGlnIleuGlnIleuGlnIleuGlnIleu 250
Db 1390 CTGGGGCTGAGACACAGCTGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1449
Qy 251 LeuThrIleAsnSerGlnProAlaValAsnGlnIleuYsSerAspSerProThrValGly 270
Db 1450 CTCACATCAACATCAACAGCCCAACATCAACAGGAGAGAGAGAGAGAGAGAGAGAG 1509
Qy 271 TTPelGlyProGlyIleTyrValTyrGlnIleuYsAlaTyrValGluPhePheCysSerIys 290
```


Db 1510 TGGGGCCCCAGGGGGCTATGCTTCCAGAAAGCCCTACTTAGAGTTTTCACCTCCCGC 1569
QY 291 GILULYSLEuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGGAAAGCACTTCTGCAAGTGTCTGAGAGACAGAGCTCCGGGTTAATTATTCAC 1629
QY 310 AlaValAsnLysAspGlySerTyrPlySerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTGTGCAATGTGAAGGGTGAAGAAACATCCCAATGCCCTGACAGCCGCGGATGTGTGC 1689
QY 330 ThrTrpGlyValAlaPheProAlaLysGluLeuIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGGCATCTTCCCTGGGGAGAGATCATCCAGCCCAACCGTAGAGATCCCGTCAGC 1749
QY 350 PheAsnValTrpLysAspGluAlaPheGluLeuIleTyrPheSerArgLysTrpAlaSerLeuTyr 369
Db 1750 TTCAATGTTCTGGAAGAGACAGGCTTTCCTGTGGATGTAGCGGCTGGGAAGACTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db 1810 GAGAGAGAGACTCCCGCTCCCGCACCATCATCCAGTACATCCAGACGACAACTACTCTGTGC 1869
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGTGACATGACTTCCCACTGGACAACTGCTTGGCAGGTGTGGAGAGAC 1926

RESULT 5
US-09-728-910-1
; Sequence 1, Application US/09728910
; Patent No. US20010025030A1
; GENERAL INFORMATION:
; APPLICANT: Rozen, Rima
; APPLICANT: Sekhon, Rajpreet
; TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETERAHYDROFOLATE
; TITLE OF INVENTION: REDUCTASE AND USES THEREOF
; FILE REFERENCE: 04844/006001
; CURRENT APPLICATION NUMBER: US/09/728,910
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/258,928
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1980)
US-09-728-910-1

Alignment Scores:
Pred. No.: 3,09e-104 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Gaps: 14
DB: 10 Gaps: 7

US-09-720-451-6 (1-408) x US-09-728-910-1 (1-2220)

QY 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
Db 676 GCCGATTTCATCATCATCGAAGCTTTTCTTGAGCTGACACATCTTCGCCGTTTGTAAG 735
QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
Db 736 GCATGCAACCGACATGGGCTCACTTGCACCACTGCCGGAATCTTCCCATCCAGAGGC 795
QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
Db 796 TACCACTCCCTTGGCAGCTGTGGAAGCTGTCCAAAGCTGAGAGTCCACAGAGATCAAG 855

QY 61 AlaAlaLeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
Db 856 GAGCTATTGAGCCATCAAGACAAACAGATGCTGCATCCGCACTAGCATGAGACTG 915
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr 99
Db 916 GCCGTGAGCTGTGCCAGAGAGCTTGTGGCCAGTGGCTTGGGCCAGAGCTTCCACTTCC 975
QY 100 ThrIleAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
Db 976 ACCCTCAACCGCAGATGCTTACACAGAGTGTCAAGCCCTGGGATGTGACTGAG 1035
QY 120 SerLysValSerArgSerLeuProTTrpArgArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCCC---AGGGCTCCCTACCTTGGGCTCTCACTGCCCCACCCCAAGCCCGAGAGAA 1092
QY 140 AspValArgProIlePheThrPheAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTACGTCCTCATCTTGTGGGCTCTCCAGCAACAAAGATACATCTTACCGTACCCAGAG 1152
QY 160 TrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1153 TGGAGCAGATTCCCTTAACGCGCTGGGCAATTCCTCTCCCTGCTTGGGGAGACTG 1212
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLeuValGluGlu 197
Db 1213 AAGCACTACTACCTTCTTACCTGGAAGAGCAAGTCCCAAGAGAGAGACTGCTAAGATG 1272
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGAGAGCTGACAGTGAACAGCAAGTGTGGAAGCTTGTCTTACCTCTCG 1332
QY 218 GlyLysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
Db 1333 GGAACACCAACCGGATGGTCAAAAGTCACTTCCCTCCCTGGAGAGAT---GAGCC 1389
QY 221 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGCGCTAGACCAAGCCTGTGGAAGAGAGATGTGGGGGTGAACGCCGAGGATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValAlaAsnGlyLysSerAspSerProThrValGly 270
Db 1450 CTCACATCAACTACACAGCCCAATCAACAGGAGAGCGTCTCCGACCCCATGCTGAGGC 1509
QY 271 TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 1510 TGGGGCCCAAGGGGGCTATGCTTCCAGAAAGGCTTACTTACCTTCCCGC 1569
QY 291 GILULYSLEuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGGAAAGCACTTCTGCAAGTGTCTGAGAGACAGAGCTCCGGGTTAATTATTCAC 1629
QY 310 AlaValAsnLysAspGlySerTyrPlySerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTGTGCAATGTGAAGGGTGAAGAAACATCCCAATGCCCTGACAGCCGCGGATGTGTGC 1689
QY 330 ThrTrpGlyValAlaPheProAlaLysGluLeuIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGGCATCTTCCCTGGGGAGAGATCATCCAGCCCAACCGTAGAGATCCCGTCAGC 1749
QY 350 PheAsnValTrpLysAspGluAlaPheGluLeuIleTyrPheSerArgLysTrpAlaSerLeuTyr 369
Db 1750 TTCAATGTTCTGGAAGAGACAGGCTTTCCTGTGGATGTAGCGGCTGGGAAGACTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db 1810 GAGAGAGAGTCCCGCTCCCGCACCATCATCCAGTACATCCAGACGAACTACTCTGTGC 1869
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGTGACATGACTTCCCACTGGACAACTGCTTGGCAGGTGTGGAGAGAC 1926

RESULT 6

| | | |
|---|-----------------------------|---|
| | US-09-770-445-690/C | : Sequence 690, Application US/09/770445 |
| | Patent No. US2002003281A1 | |
| | GENERAL INFORMATION: | |
| / | APPLICANT: | Goriach, Jörn |
| / | APPLICANT: | An, Yong-Qiang |
| / | APPLICANT: | Hamilton, Carol M. |
| / | APPLICANT: | Priest, Jennifer L. |
| / | APPLICANT: | Raines, Tracy M. |
| / | APPLICANT: | Yu, Yang |
| / | APPLICANT: | Ramaka, Joshua G. |
| / | APPLICANT: | Page, Amy |
| / | APPLICANT: | Mathew, Abraham V. |
| / | APPLICANT: | Ledford, Brooke L. |
| / | APPLICANT: | Wessner, Jeffrey P. |
| / | APPLICANT: | Haas, William David |
| / | APPLICANT: | Garcia, Carlos A. |
| / | APPLICANT: | Kricker, Maja |
| / | APPLICANT: | Slader, Red |
| / | APPLICANT: | Davis, Keith R. |
| / | APPLICANT: | Allen, Keith |
| / | APPLICANT: | Hoffman, Neil |
| / | APPLICANT: | Hurban, Patrick |
| / | TITLE OF INVENTION: | Expressed Sequences of Arabidopsis |
| / | TITLE OF INVENTION: | Thaliana |
| / | FILE REFERENCE: | 2023US (PARA-012PRV) |
| / | CURRENT APPLICATION NUMBER: | US/09/770,445 |
| / | CURRENT FILING DATE: | 2001-01-26 |
| / | PRIOR APPLICATION NUMBER: | US 60/178,472 |
| / | PRIOR FILING DATE: | 2000-01-27 |
| / | NUMBER OF SEQ ID NOS: | 999 |
| / | SOFTWARE: | FastSeq for Windows Version 4.0 |
| / | SEQ ID NO 690 | |
| / | LENGTH: | 832 |
| / | TYPE: | DNA |
| / | ORGANISM: | Arabidopsis thaliana |
| / | US-09-770-445-690 | |
| | | |
| Alignment Scores: | | |
| Pred. NO.: | 1.64e-101 | Length: 832 |
| Score: | 950.00 | Matches: 174 |
| Percent Similarity: | 87.84% | Conservative: 21 |
| Best Local Similarity: | 78.38% | Mismatches: 27 |
| Query Match: | 43.44% | Indels: 0 |
| DB: | 10 | Gaps: 0 |
| | | |
| US-09-720-451-6 (1-408) x US-09-770-445-690 (1-832) | | |
| OY | 186 | AirProArAlaIarAspLysLysLeuValGluGlnTrpAlaValProLeuLysSerVal 205 |
| Db | 830 | CGTCCGCCGAGCAGCGACAAAGACTTCACCAAGAATGGGTGTCCCTGAAAAGTT 77U |
| OY | 206 | GIASpILlTYqYLARPhaRgIArPheRgIArLeuTYCYSLSeuGLYrLysleuArSerAsnProtRp 225 |
| Db | 770 | GAGGARATTTCAGAGAAATTCACAGAGACTTGCTGCTGGGAACCTTAATAAACAGAGTCCCTGG 711U |
| OY | 226 | SerGluLeuAspLYLeuGlnProGlnTrpLysILElleasnGluGlnLeuGlnLysILE 245 |
| Db | 710 | TCTGAATTAGAGTGCAGCTCCAGCCAGACCAAGGATTAATAAGCAGCACCTGATTAAGATC 651 |
| OY | 246 | AsnThrLysGLYPheLeuThrILLEasnSerGlnProAlaValAsnGLYgluLysSerAsp 265 |
| Db | 650 | AACGCCAAAGGCTTCTTGACCATCAATAGCCAAACATCAAGTCAAGCCGAGAGATCTGAT 591 |
| OY | 266 | SerProThrValGLYTrpGLYGLYProGLYGLYTY-eValTYrGlnLysAlaTY-eValGlu 285 |
| Db | 590 | TCCCACAAGCTTGATGGGAGAGTCTGTTGGATGATGTATACCANAAGGCTTAATCTAGAG 531U |
| OY | 286 | PhaPheCySerLysGLYLeuLysLeuAspALLEValAspLysCYSLysAspArgThrSer 305 |
| Db | 530 | TTCCTTCGCTCAAAGGAGAAATTAATGACGTGAGTGGAGAAATGCAAAAGCTTTGCCATCG 471U |
| OY | 306 | IeuThrTYMetAlaValAsnLysAspLYSerTrpLysSerAsnValGLYGLNThrAsp 325 |

```

Db      470  ATTTACTTACATGCTCTGTGAACAAGAGACAGTGGGATATCGAACAACACTGCCCAAGCTGAT 411
               ::::::::::::::::::::::: ||| ||||| ||| |||
QY      326  ValAsnAlaValThrTrpGlyValPheProAlaLysGluLeileglnProThrIleVal 345
               |||:::|||||:::||||| ||| ||| ::|::| |||
Db      410  GTGATATCTTAACTTGGGGAGTTTCCCGCTAAGGAATCATCAACCAACCATTTGTC 351
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      346  AspProValSerPheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgLys 365
               |||:::|||||:::||||| ||| ||| ||| ||| |||
Db      350  GATCCCGCTAGCTTCAACGCTCGAAGGATGAACATTTGGACTTGGTCACGAAGCTGG 291
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      366  AlaSerLeuTrpProGluAspGluAlaSerArgLysLeuGluGluValGlyLys 385
               |||:::|||||:::||||| ||| ||| ||| ||| |||
Db      290  GCTACTTGTACCAAGACCAACCCCTTCAGAAACTTGTCTCGAGGATGGAAGAACAGC 231
               |||:::|||||:::||||| ||| ||| ||| ||| |||
Db      230  TACTATTGGTGAAGCTTGTGGAGAACGATACATCAATGTGTACATATTCGGGCTCTT 171
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      406  AlaAsp 407
               |||||
Db      170  GCTGAT 165

RESULT 7
US-10-113-852A-1
; Sequence 1, Application US//10113852A
; Publication No. US20020192784A1
; GENERAL INFORMATION:
; APPLICANT: Appling, Dean R.
; APPLICANT: Hanson, Andrew D.
; APPLICANT: Raymond, Rhonda R.
; APPLICANT: Kojer, Sanja
; TITLE OF INVENTION: Biosynthesis of S-Adenosyl Methionine in a Recombinant Yeast S
; FILE REFERENCE: 119927-1033
; CURRENT APPLICATION NUMBER: US/10/113, 852A
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/280,333
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (826)..(2625)
; OTHER INFORMATION: DNA construct
US-10-113-852A-1

Alignment Scores:
Pred. NO.: 1.43e-80 Length: 3082
Score: 778.00 Matches: 168
Percent Similarity: 59.37% Conservative: 76
Best Local Similarity: 40.88% Mismatches: 147
Query Match: 35.57% Indels: 20
                Gaps: 11

US-09-720-451-6 (1-408) x US-10-113-852A-1 (1-3082)
QY      4  IleValThrGluLeuPheThrAspThrAspIlePheLeuLysPheValAsnAspCysArg 23
               |||:::|||||:::||||| ||| ||| ::|::| |||
Db      1378  ATCATCATCATGATGTTTATTCATGTGTATTAATTACCACTACCTGTTTCCCAAGTTAA 1437
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      24  GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsnYrrLysGly 43
               |||:::|||||:::||||| ||| ||| ||| ||| |||
Db      1438  GCTCGCGGACATGACGTCGCCATTAATTCGCCGATCATGCGATCACTACTACGCGGCGC 1497
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      44  PheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMetAlaIleVal 63
               |||:::|||||:::||||| ||| ||| ||| ||| |||
Db      1498  TTCTTGGAGAAGATCCAAATGAGGCGCAAAATTCATCTCTCAACATTTTCGTCCGATTG 1557
               |||:::|||||:::||||| ||| ||| ||| ||| |||
QY      64  GluProIleLysAspAsnGluGluAlaValLysAlaTrpGlyIleHisLeuGlyThrGlu 83
               |||:::|||||:::||||| ||| ||| ||| ||| |||

```



```
Db 427 CCAATTTCGGGCAACGACCTAGAGCTACATTTCACGAAACCAACGCTGGACGAT 486
QY 163 TYPProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyr 182
Db 487 TTTCCTCGTGGAAGGTGGGAGATAGGCC-AACCTGCTTACGGCAGCTTCAGCGATCAT 545
QY 183 GlnPheMetArgProArg 188
Db 546 CAGTTCACGACGAGAAG 563

RESULT 9
US-09-878-574-3765
; Sequence 3765, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3765
; LENGTH: 313
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-006-Q1-B1-C9
US-09-878-574-3765

Alignment Scores:
Pred. No.: 8,89e-38 Length: 313
Score: 398.00 Matches: 78
Percent Similarity: 94.12% Conservative: 2
Best Local Similarity: 91.76% Mismatches: 5
Query Match: 18,208 Indels: 0
Gaps: 0

US-09-720-451-6 (1-408) x US-09-878-574-3765 (1-313)
QY 322 GYGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLysGlnIleIleGln 341
Db 1 GGCACAACTGACGTTAATGCTGTGACATGGGGGTCTTCCACGCTAAGAGATAATCTCA 60
QY 342 ProThrIleValAspProValSerPheAsnValTrpLysAspGluAlaPheGlnIleTrp 361
Db 61 CCACCATGTGTAGATCCTGTGACGCTTCATGTATGAGAGATGACCATTTGAATTTGG 120
QY 362 SerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGluGln 381
Db 121 TCAAGATGATCGGCAAGCTGTGACCTCGAGGATTAAGCATCCAGAAATGTGTGAAGAG 180
QY 382 ValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyPheLeu 401
Db 181 GTTGGGGGACGACCTTCTTGGTTACTTGGTGCACAAATTATTAATCATGATGATCCT 240
QY 402 PheAlaValAlaPheAla 406
Db 241 TTGCGCGTCTTGCA 255

RESULT 10
US-09-923-876-5596
; Sequence 5596, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laljundi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
```

```
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5596
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457310H1
; LOCATION: 81, 183
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5596

Alignment Scores:
Pred. No.: 5,86e-37 Length: 264
Score: 390.00 Matches: 72
Percent Similarity: 87.50% Conservative: 5
Best Local Similarity: 81.82% Mismatches: 10
Query Match: 17,83% Indels: 1
Gaps: 0

US-09-720-451-6 (1-408) x US-09-923-876-5596 (1-264)
QY 116 LeuIleGlnGlnSerLysValSerArgSerLeuProTrpAlaArgProAlaAsnValPhe 135
Db 3 TTAATGTAGAGATCCAAAGCTTCAAGGCATTAACCTTGAGGCCAGGACGACTAAGTTTC 62
QY 136 ArgValLysGluAspVal-ArgProIlePheTrpAlaAsnArgProLysSerTyrIleSe 155
Db 63 CGGTGTAAGAGATGTTNCACACTATATTTCTGGGCCAACAGACCAAGACTATCTTAA 122
QY 155 ArgThrIleGlyTrpAspGlnTyrProHisGlyArgTrpLysAspSerCysAsnProSe 175
Db 123 AAGGACATTAGATTGGGATGATAGTATCCCATGACGAGGCGGTGATTCGTGAAACCATC 182
QY 175 rTyrGlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLeuVal 195
Db 183 NCATGACACTTACTGTGACACGATTCACAAAGCACGAGCCGTGTAAAGACTCA 242
QY 195 lGlnGlnTrpAlaValProLeu 202
Db 243 AGAGAAATGGCTGTTCACATG 264

RESULT 11
US-09-923-876-476
; Sequence 476, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laljundi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 476
; LENGTH: 265
; TYPE: DNA
; ORGANISM: Zea mays
```

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157863.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 54.1%; Score 1184; DB 21; Length 497;
Best Local Similarity 78.9%; Pred. No. 9.7e-110;
Matches 213; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 ARVIYTOLEFDIFLKRVNDCKRGITGTCPIVPGIMPNYKGRITMGFCCKTKIPADIM 60
DB 214 ADLIVTOLEFDIFLKRVNDCKRGITGTCPIVPGIMPNYKGRITMGFCCKTKIPADIM 273

QY 61 ALEPIKNDNEAVKAYGIHLGTEMCKKILAHGKITLHLYTLNMKSALATIMNLGIEES 120
DB 274 ALEPIKNDNEAVKAYGIHLGTEMCKKILAHGKITLHLYTLNMKSALATIMNLGIEES 333
QY 121 KVSRLPWRPAPVFRKEDVRPIFWANRPKSYISRTIGMDQYPHRGWDSNPSYGALS 180
DB 334 KVSRLPWRPAPVFRKEDVRPIFWANRPKSYISRTIGMDQYPHRGWDSNPSYGALS 393
QY 181 DYQFMRPRADKLVEMAVPLKSVEDYIFRFLYCLGKIRSPMSLDGLOPETKINE 240
DB 394 DYQFMRPRADKLVEMAVPLKSVEDYIFRFLYCLGKIRSPMSLDGLOPETKINE 453
QY 241 QLEKINTKGFLLTINSOPAVNGEKSDSPYV 270
DB 454 QLEKINTKGFLLTINSOPAVNGEKSDSPYV 483

RESULT 14
AAR8358 standard; Protein; 656 AA.
ID AAR8358
XX AAR8358;
AC AAR8358;
XX 15-OCT-1995 (first entry)
DT 15-OCT-1995 (first entry)
XX Human methylene-tetrahydrofolate-reductase.
DE Human methylene-tetrahydrofolate-reductase.
XX Methylenetetrahydrofolate reductase; MTHFR; gene therapy.
KW cadiocvascular disease; neurological disease; folic acid metabolism;
KW EC-1.5.1.20; enzyme.
XX Homo sapiens.
OS Homo sapiens.
XX W09533054-A1.
PN W09533054-A1.
XX 07-DEC-1995.
PD 07-DEC-1995.
XX 25-MAY-1995; 95WO-CA00314.
PF 25-MAY-1995; 95WO-CA00314.
XX 26-MAY-1994; 94GB-0010620.
PR (UYMC-) UNIV MCGILL.
PA (UYMC-) UNIV MCGILL.
XX Goyette P, Rozen R;
PI Goyette P, Rozen R;
DR WPT: 1996-030565/03.
DR N-PSDB; AAT09689.
DR Human methylene:tetra:hydro:folate reductase cDNA probe - for
PT detection of sequence abnormalities in methylene:tetra:hydro:folate
PT reductase e.g. in cardiovascular, neurological or folic acid
PT metabolism disorders
XX Dislosure: Fig. 6A-6C; 66pp; English.
PS This is the protein sequence encoding human MTHFR, the gene of
CC which has been localised to chromosome 1p36.3. Deficiencies of
CC this protein may lead to cardiovascular and neurological disorders
CC and disorders influences by folic acid metabolism.
XX Sequence 656 AA;
SQ

Query Match 44.8%; Score 979; DB 17; Length 656;
Best Local Similarity 44.9%; Pred. No. 6e-89;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;

QY 1 ARVIYTOLEFDIFLKRVNDCKRGITGTCPIVPGIMPNYKGRITMGFCCKTKIPADIM 60
DB 222 ADLIVTOLEFDIFLKRVNDCKRGITGTCPIVPGIMPNYKGRITMGFCCKTKIPADIM 281
QY 61 ALEPIKNDNEAVKAYGIHLGTEMCKKILAHGKITLHLYTLNMKSALATIMNLGIEES 119
DB 214 ADLIVTOLEFDIFLKRVNDCKRGITGTCPIVPGIMPNYKGRITMGFCCKTKIPADIM 273

DB 282 DVIEPIKNDAAIRNYGIELAVSLCQELLASGLVPLGHEFTLNRREMATTEVLKRLGMWTE 341
 QY 120 SKVSESLPWRPANKYVKEVDRIFWANRPKSYISRTIGWDQYPHGRWGDCSNPSYCAL 179
 DB 342 DP-RRLPWLALSAHPKREEDVRPIFWASRPKSYIYRTQEWDEFFNGRWGNSSSPAFAGEL 400
 QY 180 SDYQ--FMRPRARDKLVVEAVNPLKSVEDIYERFRLYCLGKLRN-----PWSELDG 230
 DB 401 KDYLFLYKLSKSPKEELLKMWGEELTSEASVFEVLYLSGEPNRRNGHKVTCLPWND-EP 459
 QY 231 LQPEKTIINEOLEKINTKGFLTINSQPAVNGEKSDPTVWGSGPGGYVQKAYVEFFCSK 290
 DB 460 LAAETSLKEELLRVNRQGITLINSQPNINGKPSDPTVWGSGPGGYVQKAYLEFFTSR 519
 QY 291 EKLDALVKCKD-RFSLTYMAVNRKGSWKSNGVGTDNVAVTGWVFPKAEIIQPTIVDPVS 349
 DB 520 ETAEALLQVLLKYLRYNHLVNVKGENITNAPELOPNVATWGIFFGREIIQPTIVDPVS 579
 QY 350 FNVKDEAFETWSRGWASLPEDEASRKLVEEYGGSHFLVSLVDNDY-INGDLFAVFAD 407
 DB 580 FMFWKDEAFALWIERWGKLYEESPSRTIIQYIHDNYFLVNLVNDNDFPLDNCMLQWVED 638

RESULT 15
 AAEL2607
 ID AAEL2607 standard; Protein; 656 AA.
 XX
 AC AAEL2607;
 XX
 XX 03-JAN-2002 (first entry)
 DE Human methylenetetrahydrofolate reductase (MTHFR) protein #2.
 XX Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
 KW antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
 KW pancreas; lung; brain; cystostatic; colon carcinoma; breast; colorectal;
 KW neuroblastoma; leukaemia.
 OS Homo sapiens.
 XX
 XX US2001025030-A1.
 XX
 XX 27-SEP-2001.
 XX
 XX 01-DEC-2000; 2000US-0728910.
 XX
 XX 01-MAR-1999; 99US-0258928.
 XX
 XX (ROZE/) ROZEN R.
 XX (SEKH/) SEKHON J.
 XX
 XX Rozen R, Sekhon J;
 XX
 XX WPI; 2001-638509/73.
 XX
 XX N-PSDB; AAD20463.
 XX
 XX New antisense nucleic acids, which are methylenetetrahydrofolate
 PT reductase inhibitors, useful for treating, stabilising or preventing
 PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
 PT neuroblastoma
 XX
 XX Disclosure; Fig 6; 68pp; English.
 XX
 XX The invention relates to a non allele-specific antisense nucleic acids,
 CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
 CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
 CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
 CC 5-methylenetetrahydrofolate, a co-substrate for methylation of
 CC homocysteine to methionine. The invention provides potential therapy for
 CC individuals with MTHFR deficiency. The non allele-specific antisense
 CC nucleic acids are useful for treating, stabilising or preventing cancer,
 CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
 CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,

CC neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
 CC therapy. The present sequence is human methylenetetrahydrofolate
 CC reductase (MTHFR) protein #2. Human MTHFR gene is mapped to
 CC chromosome 1p36.3.
 XX
 SQ Sequence 656 AA;
 Query Match 44.8%; Score 979; DB 22; Length 656;
 Best Local Similarity 44.9%; Pred. No. 6e-89;
 Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;
 QY 1 ARVITVQLFYTDLFLFVNDRCQIGITCPVPGIMPINNYKGFIRMTCFKTKIPADIM 60
 DB 222 ADFTITQFFADFFRFVAKCTDMGITCPVPGIPIQGIQHSLSQLVKLSKLEVPQEI 281
 QY 61 AALEPIKNDNEAVKAYGIHLGTEMCKKILAHG-IKTLHLTYLNMKESALAIMNLGLIEE 119
 DB 282 DVIEPIKNDAAIRNYGIELAVSLCQELLASGLVPLGHEFTLNRREMATTEVLKRLGMWTE 341
 QY 120 SKVSESLPWRPANKYVKEVDRIFWANRPKSYISRTIGWDQYPHGRWGDCSNPSYCAL 179
 DB 342 DP-RRLPWLALSAHPKREEDVRPIFWASRPKSYIYRTQEWDEFFNGRWGNSSSPAFAGEL 400
 QY 180 SDYQ--FMRPRARDKLVVEAVNPLKSVEDIYERFRLYCLGKLRN-----PWSELDG 230
 DB 401 KDYLFLYKLSKSPKEELLKMWGEELTSEASVFEVLYLSGEPNRRNGHKVTCLPWND-EP 459
 QY 231 LQPEKTIINEOLEKINTKGFLTINSQPAVNGEKSDPTVWGSGPGGYVQKAYVEFFCSK 290
 DB 460 LAAETSLKEELLRVNRQGITLINSQPNINGKPSDPTVWGSGPGGYVQKAYLEFFTSR 519
 QY 291 EKLDALVKCKD-RFSLTYMAVNRKGSWKSNGVGTDNVAVTGWVFPKAEIIQPTIVDPVS 349
 DB 520 ETAEALLQVLLKYLRYNHLVNVKGENITNAPELOPNVATWGIFFGREIIQPTIVDPVS 579
 QY 350 FNVKDEAFETWSRGWASLPEDEASRKLVEEYGGSHFLVSLVDNDY-INGDLFAVFAD 407
 DB 580 FMFWKDEAFALWIERWGKLYEESPSRTIIQYIHDNYFLVNLVNDNDFPLDNCMLQWVED 638

Search completed: February 14, 2003, 23:07:42
 Job time : 73 secs

```

; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700157001H1
US-09-923-876-476

```

Alignment Scores:

```

Pred. No.: 7 71e-37      Length: 265
Score: 389.00      Matches: 71
Percent Similarity: 92.86%      Conservative: 7
Best Local Similarity: 84.52%      Mismatches: 6
Query Match: 17.79%      Indels: 0
DB: 10      Gaps: 0

```

US-09-720-451-6 (1-408) x US-09-923-876-476 (1-265)

```

QY 218 GILYSLNLAAGSERAPROTHRPSERGLNLEUASAPGLYENGINPROGLNTHRLYSILE 237
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 2 GGGAACTCTCAAGACAGCCATGCTCAAGATTGGAGGCTCTTCAACGAGCAAGAAATT 61
QY 238 TLEASNGIUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNU 257
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 62 ATGATGAGCCAGTGTGGTGAATATTAACCAAGAGGTTTCTTACAAATTACAGCAACCT 121
QY 258 ALAVALASNGIUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNU 277
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 122 GCTGTAAATGGAGAGAAATCCAGCTCGCTACTGTGTGGTGGTGGTGGTGGTGGTGGTGG 181
QY 278 VALTYGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGL 297
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 182 GTTATACAGAGAGCCATCTCTCAATTTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 241
QY 298 ASPLYSCTSLYS 301
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 242 GAGAAATCAAA 253

```

RESULT 12

```

US-09-923-876-5753
; Sequence 5753, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5753
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457642H1
; LOCATION: 2, 165, 173, 180
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5753

```

Alignment Scores:

```

Pred. No.: 7.58e-25      Length: 181
Score: 284.00      Matches: 50
Percent Similarity: 88.14%      Conservative: 2
Best Local Similarity: 84.75%      Mismatches: 7
Query Match: 12.99%      Indels: 0
DB: 10      Gaps: 0

```

US-09-720-451-6 (1-408) x US-09-923-876-5753 (1-181)

```

QY 116 LEUILLGLUGLUSERLYSVALSERAPROTHRPSERGLNLEUASAPGLYENGINPROGLN 135
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 3 TTATATGAGAGAGTCCAGAGTTCACAGCCATTTACTTGGAGGCCAGAGCAATATGTTTTC 62
QY 136 ARGVALYSLGLUAPVALARGPROLLEPHERPALAASARGPOLYSSERTYRILLESER 155
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 63 CGGTGTAAGAGAGAGAGTTCGACCTTATTTCTGGGCAACAGACCAAGAGAGAGAGAGAGAG 122
QY 156 ARGTHRLLEGLYTRPASPGLINRPROHISGLYARTRPGLYASPSERSCYASAPRO 174
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 123 AGGACATTAAGTGTGGAGATACGATCCCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 179

```

RESULT 13

```

US-09-923-876-1503
; Sequence 1503, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Itc)
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 1503
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700158929H1
US-09-923-876-1503

```

Alignment Scores:

```

Pred. No.: 1.46e-23      Length: 216
Score: 274.00      Matches: 51
Percent Similarity: 81.69%      Conservative: 7
Best Local Similarity: 71.83%      Mismatches: 13
Query Match: 12.53%      Indels: 0
DB: 10      Gaps: 0

```

US-09-720-451-6 (1-408) x US-09-923-876-1503 (1-216)

```

QY 274 PROGLYGLYTYRYVALTYRGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNUGLNU 293
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1 CCGTAGGAGCTTATATCAGAGAGCCATCTCAATTTCTGCGAAGAGAGAGAGAGAGAGAG 60
QY 294 ASPALALEUVALASPLYSCTSLYSASPARGTHSERLEUTHRYMETALALAASNYLS 313
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 61 GACCACTATATGAGAGATCAAGACATTCCTCTCTCACTTACATGCTGTGAACAG 120
QY 314 ASPGLYSERTPLYSERSTANALGLYLNTHRSVALASNAALVALNTHRTTPGLYVAL 333
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 121 GATGAGAGAAACATTTCTCAATATTTCCACCAAGAGCCGTAATGCTGTCAAGTGGGTGTT 180
QY 334 PHEPROALALYSLNULEILEGINPROTHRLLE 344
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 181 TTCCCTGCAAGAGAGATATTCAGAGCCCTACGGTG 213

```

RESULT 14

```

US-09-294-093B-4653
; Sequence 4653, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.

```

```
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4653
; LENGTH: 159
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO. US20010051335A1 700354741H1
; NAME/KEY: unsure
; LOCATION: 14
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4653

Alignment Scores:
Pred. No.: 1.5e-21 Length: 159
Score: 255.00 Matches: 43
Percent Similarity: 94.00% Conservative: 4
Best Local Similarity: 86.00% Mismatches: 3
Query Match: 11.66% Indels: 0
DB: 10 Gaps: 0

US-09-720-451-6 (1-408) x US-09-294-093B-4653 (1-159)
QY 133 AsnValphearglyalysgluaspvalargprollephetpallaasnaargprolyser 152
Db 8 AATGTTTCGGTAAAGAGATGTGACCTATATCTGGCCACAGACCAAGAGC 67
QY 153 TyrlleserAargThrllleGIYTPAspGlnTYrProHlsGlyAArgTrpGlyAspSerCys 172
Db 68 TATCTTAAAGACATTAGTGGTGGATCATCCCATGACGCGTGGGTGATCTCGG 127
QY 173 AsnProserTYrGlyAlaLeuSerAspTYr 182
Db 128 AACCCATCATATGAGACACTTACTGACAC 157

RESULT 15
US-09-734-017A-63
; Sequence 63, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Clippus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides and
; TITLE OF INVENTION: nucleosides
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/771,100
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 63
; LENGTH: 409
; TYPE: DNA
```

```
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(409)
; OTHER INFORMATION: 62_mm20_c10rev
US-09-734-017A-63

Alignment Scores:
Pred. No.: 5.1e-16 Length: 409
Score: 213.00 Matches: 42
Percent Similarity: 89.80% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 4
Query Match: 9.74% Indels: 1
DB: 10 Gaps: 0

US-09-720-451-6 (1-408) x US-09-734-017A-63 (1-409)
QY 3 ValIleValThrGlnLeuPheTYrAspThrAspIlePheLeuYspPheValAsnAspCys 22
Db 260 GTCATATACACCCAGCAGCTGTTTATGATACCATATCTTTTGAATTTGGAATGATTCGT 319
QY 23 ArgGlnIleGlyIleThcYsProIleValProGlyIleMetProIleAsnAsnTYrIlys 42
Db 320 CGTCAATTGGATCAAGGTGCCATTGTACTGTGATCATGCCATCAAAATTACAG 379
QY 43 -GlyPheIleArgMetThrGlyPhe 50
Db 380 GGGCTTCTCCGCAATGACCACCTTY 404
```

Search completed: February 12, 2003, 17:07:24
Job time : 93 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 14, 2003, 23:07:52 : Search time 3547 Seconds
(without alignments)
3347.602 Million cell updates/sec

Title: US-09-720-451-6
Perfect score: 2187
Sequence: 1 ARVITQLEFYDTDFLKRNV.....VSLVDNDYINGDLFANFADF 408

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+P2n.model -DEV-X1h
-O/-cgn2_1/USFPO.spool/US09720451/runat_05022003_074814_2510/app_query.fasta_1.583
-DB=GenDb1 -GFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=Dits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720451.cgn2_1.1616_@runat_05022003_074814_2510 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -YGAPOP=10 -XGAPOP=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenDb1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_st:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_st:*
28: em_un:*

29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|--------------|---------------------|
| 1 | 1834 | 83.9 | 1897 | 8 | AF181967 | AF181967 Arabidops |
| 2 | 1834 | 83.9 | 1917 | 8 | AF050434 | AF050434 Arabidops |
| 3 | 1834 | 83.9 | 2034 | 8 | AF370515 | AF370515 Arabidops |
| 4 | 1831 | 83.7 | 2037 | 8 | AF1245414 | AJ245414 Arabidops |
| 5 | 1777 | 81.3 | 1810 | 8 | AY122922 | AY122922 Arabidops |
| 6 | 1777 | 81.3 | 1994 | 8 | AF181966 | AF181966 Arabidops |
| 7 | 1777 | 81.3 | 2002 | 8 | AF070034 | AF070034 Arabidops |
| 8 | 1762 | 80.6 | 2225 | 8 | AF174486 | AF174486 zea mays |
| 9 | 1522 | 69.6 | 109741 | 8 | AC004005 | AC004005 Arabidops |
| 10 | 1424 | 65.1 | 100906 | 8 | AF244516 | AF244516 Arabidops |
| 11 | 1411 | 64.5 | 1785 | 12 | AR441241 | AR441241 Synthetic |
| 12 | 979 | 44.8 | 1971 | 6 | AX069359 | AX069359 Sequence |
| 13 | 979 | 44.8 | 2187 | 6 | AR144956 | AR144956 Sequence |
| 14 | 979 | 44.8 | 2187 | 6 | AR144956 | AR144956 Sequence |
| 15 | 979 | 44.8 | 2187 | 6 | AX050441 | AX050441 Sequence |
| 16 | 979 | 44.8 | 2196 | 12 | HSU09806 | U09806 Synthetic |
| 17 | 979 | 44.8 | 2219 | 6 | A47328 | A47328 Sequence 3 |
| 18 | 979 | 44.8 | 2219 | 6 | AR097969 | AR097969 Sequence |
| 19 | 979 | 44.8 | 2219 | 6 | AR145811 | AR145811 Sequence |
| 20 | 979 | 44.8 | 2219 | 6 | AX356174 | AX356174 Sequence |
| 21 | 979 | 44.8 | 2219 | 6 | AX452351 | AX452351 Sequence |
| 22 | 979 | 44.8 | 2220 | 6 | A47328 | A47328 Sequence 1 |
| 23 | 979 | 44.8 | 2220 | 6 | AR097968 | AR097968 Sequence |
| 24 | 979 | 44.8 | 2220 | 6 | AR145810 | AR145810 Sequence |
| 25 | 979 | 44.8 | 2220 | 6 | AX356172 | AX356172 Sequence |
| 26 | 979 | 44.8 | 2220 | 6 | HS4237672 | AJ237672 Homo sapi |
| 27 | 815.5 | 37.3 | 38571 | 3 | SPAC56F8 | U59849 Caenorhabdi |
| 28 | 812 | 37.1 | 81624 | 3 | U39849 | U39849 Caenorhabdi |
| 29 | 778 | 35.6 | 3082 | 8 | AC099763 | AC099763 Caenorhab |
| 30 | 778 | 35.6 | 3082 | 8 | SCYGL125W | Z72647 S. cerevisia |
| 31 | 778 | 35.6 | 3082 | 8 | SCSUP44GN | X94106 S. cerevisia |
| 32 | 629 | 28.8 | 3255 | 8 | SPAC343 | AF011686 Schizosac |
| 33 | 578.5 | 26.5 | 42947 | 8 | SPAC343 | AF011686 Schizosac |
| 34 | 452 | 20.7 | 1250 | 10 | RNDU57049 | U57049 Rattus norv |
| 35 | 429.5 | 19.6 | 198728 | 8 | AL606929 | AL606929 Mouse DNA |
| 36 | 421.5 | 19.3 | 1746 | 10 | D89118 | DB9118 Schizosacch |
| 37 | 415 | 19.0 | 159273 | 2 | AC025001 | AC025001 Rattus no |
| 38 | 404.5 | 18.5 | 143039 | 2 | AC125869 | AC125869 Rattus no |
| 39 | 335.5 | 15.3 | 181581 | 2 | AC113917 | AC113917 Rattus no |
| 40 | 294 | 13.4 | 110000 | 2 | LMFLCHR36.26 | Continuation (27 o |
| 41 | 270 | 12.3 | 110000 | 2 | LMFLCHR36.02 | Continuation (3 of |
| 42 | 251.5 | 11.5 | 215050 | 1 | AL646057 | AL646057 Ralstonia |
| 43 | 250.5 | 11.5 | 1376 | 9 | HS4297061 | AJ297061 Homo sapi |
| 44 | 249 | 11.4 | 326 | 11 | G71668 | G71668 A64994334FM |
| 45 | 249 | 11.4 | 343 | 11 | G70952 | G70952 A06717634FB |

RESULT 1

ALIGNMENTS

AF181967 1897 bp mRNA linear PLN 13-DEC-1999
LOCUS AF181967
DEFINITION Arabidopsis thaliana methylenetetrahydrofolate reductase MTHFR2
ACCESSION AF181967
VERSION AF181967.1 GI:5911426
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1897)
REFERENCE Roje, S., Wang, H., McNeil, S.D., Raymond, R.K., Applin, D.R.,
AUTHORS Shachar-Hill, Y., Bohnert, H.J., and Hanson, A.D.,
TITLE Isolation, characterization, and functional expression of cDNAs
encoding NADH-dependent methylenetetrahydrofolate reductase from
higher plants
JOURNAL J. Biol. Chem. 274 (51), 36089-36096 (1999)
MEDLINE 20062814
PUBMED 10593891
REFERENCE 2 (bases 1 to 1897)
AUTHORS Roje, S. and Hanson, A.D.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1999) Horticultural Sciences, University of
Florida, Hull Road, Gainesville, FL 32611-0650, USA
FEATURES
source
location/Qualifiers
1..1897
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
1..1785
/EC_number="1.5.1.20"
/function="catalyzes the reduction of
5,10-methylenetetrahydrofolate (CH2-THF) to
5-methyltetrahydrofolate (CH3-THF)"
/note="NADH-dependent"
/codon_start=1
/product="methylenetetrahydrofolate reductase MTHFR2"
/protein_id="A055788.1"
/db_xref="GI:5911427"
/translation="MKYIDKIQSLADEGKTAFSEFFPRKTEGVDNLEFEMDMVAV
GTFYEDYWGSGSTADLTLDIASRMQNVVCSMMHLCCTMPERKIDHLELTRSN
GIQVLAIRGDPHPGODPFYVGEGFDCALDLVNHRSYGYRPIITAGYPAHAPRY
IGENGLASNEAYOSDLEFKKIDACADILVLPEDPDIPKFPVNDROISGCPY
PGIMDINNYRGELRMGFCYKIPVEYMAIRPDRIPREAVAYGILNIGTCKKMLA
HKVYSILHTLNMKSKALLMLNGLMIDSKISRSRPRKPNVRFIKEDVRLTKMA
RPRKTIISRTKMGEDFQGMGDSRSASTGALSDHQSRRKARDKLDQEWVPLKSV
DIOEKFKELGCLNKSSEWSELGDLQPERRLNEGLIKVNSKDKLQDSQVVAERS
DSPVVGWGVGYOKAYLEFCSKEKIDAVEKCKALPSITTYMAVKKGSQVNSNTA
QADVNAVWGVFPKKEIIQPTIYDPAISFVWDEAFETWSRWANLYPEADPSHMLE
EVKNSYLYVSLVENDYINGDIFAVPADL"

BASE COUNT 524 A 362 C 457 G 554 T
ORIGIN

Alignment Scores:
Pred. No.: 1,34e-152 Length: 1897
Score: 1834.00 Matches: 334
Percent Similarity: 91.15% Conservative: 37
Best Local Similarity: 82.06% Mismatches: 36
Query Match: 83.86% Indels: 0
Gaps: 0

US-09-720-451-6 (1-408) x AF181967 (1-1897)

OY 1 AAlaAGValIILeValThGleuPheThrAspThrAspIlePheIleuysPheValAsn 20
Db 559 GCGATGCTCATTTGCTTACGACGCTTTCTATATATCATATATATCCCAAGTTGGAAAT 618

OY 21 AspCysArgGlnIleGlyTLeThCysProIleValProGlyIleMetProIleAsn 40
Db 619 GATTGCGCAAAATGGATAGTGTCCATGTCTCGAATATAGCTATATAAC 678

OY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrIleProIleAsnIleMet 60

Db 679 TACAGAGCGTTTTCGATGACGCTGTTTCGAAGACAAAGATACAGATGAGGTGATG 738
OY 61 AAlaIleuGluProIleAspAsnGluGluValIleValIleValIleValIleValIle 80
Db 739 GCTGCTTGAGCGCTATACAGATATACAGAACAGCTGTGAAGCTGATGATACCTT 798
OY 81 GlyThrGluMetCysLysLysIleLeuValIleGlyIleLysThrIleuAsnIleuThr 100
Db 799 GGAACAGAGATGTGAAAAAGATGTGGCTCATGATGATGATGATGATGATGATGATG 858
OY 101 LeuAsnMetGluLysSerAlaLeuValIleLeuMetAsnLeuGluIleuGluIleuSer 120
Db 859 TTCAACATGGAATAATTCGCTCTTCCAAATATGATGAATTTGGTATGATGATGATG 918
OY 121 LysValSerArgSerLeuProThrArgArgProAlaAsnValPheArgValLysGluAsp 140
Db 919 AAAATTTCGTCTTCTTACCTCGAGAGCGCCCTCAAAATGTTTCCGTACTAGGAGAT 978
OY 141 ValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTrp 160
Db 979 GTGCGCCCATTTTCTGGGCAACCGTCCAAAGAGCTACATTTCTGAAACCAAGGCGTGG 1038
OY 161 AspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
Db 1039 GAAGATTTCACAAAGCGCGGTGGGTGATTCACGCGATCTTCATATGATGCTCCTCG 1098
OY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluIleThrAlaVal 200
Db 1099 GATCATCATGCTTCACGTCGCCGAGACGCTGACAGAGAGCTTCACAAAGATGSGGTGTC 1158
OY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
Db 1159 CCACTGAATAAGCTTGAAAGATATTCAGAGAAATTCAGAGAGCTGCTGCTTGGAAACTT 1218
OY 221 ArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIleIleAsnGlu 240
Db 1219 AAAAGAGCTCCTGCTGTGAATGATGAGCTCCAGCAGCAAGCAAGATTAATAACAG 1278
OY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
Db 1279 CAACGTATAAAGCTCAATCCAAAGGCTTCTGTACATCAATGCCAACACATGACTCAAC 1338
OY 261 GlyGluLysSerAspSerProThrValGlyTrpGlyLysProGlyLysValIleGln 280
Db 1339 GCGGAGAGTGTGATCCCAAGCTGTGAGGGAGGAGTCTGTGATATGATATACCAA 1398
OY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
Db 1399 AAGGCTTATCTAGATCTTCTGCTCAAAAGAGCAAAATTAATGATGATGATGATGATG 1458
OY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLysSerAsn 320
Db 1459 AAAAGCTTGCATGATCTTACTTACTGCTGTGACAAAGAGAGAGACAGTGGTATACGA 1518
OY 321 ValGlyIleThrAspValAlaValIleValIleValIleValIleValIleValIle 340
Db 1519 ACTGCCCAAGCTGATGTAATGCTGTAACCTGGGAGTTCCTCCGCTAAGCAAAATCAAT 1578
OY 341 GlnProThrIleValAspProValSerPheAsnValIleTrpLysAspGluAlaPheGluIle 360
Db 1579 CAACCAACATGTGATCCCGCTAGCTCAACGCTCGAAGAGATGAACATTTAGACT 1638
OY 361 TrpSerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
Db 1639 TGGTCAGAGAGCTGGGCTAACTGTATACCAAGAGCCGACCTCCAGAAATGCTGCTCAG 1698
OY 381 GluValIleGlySerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
Db 1699 GAGGTGAAGAACAGCTACTATTTGTAAGCTGTGTGAGAGACAGATTAATCAATGTGTAC 1758
OY 401 LeuPheAlaValPheAlaAsp 407

db 1759 AATATGCGGTGTTCGTGAT 1779

RESULT 2

LOCUS AY050434 1917 bp mRNA linear PLN 20-AUG-2001

DEFINITION Arabidopsis thaliana At2g44160/F6E13.29 mRNA, complete cds.

VERSION AY050434

ACCESSION AY050434.1 GI:15215809

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1917)

1 Shin, P., Chen, H., Cheuk, R., Kim, C.-J., Koeseema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Hayashizaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis cDNA clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1917)

2 Shin, P., Chen, H., Cheuk, R., Kim, C.-J., Koeseema, E., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Hayashizaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submissions

JOURNAL Submitted (02-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

FEATURES

source

1. 1917

Location/Qualifiers

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAPL05-16-12(R13348)"

/note="ecotype: Columbia"

1. 1405

11. 1825

/note="putative methylentetrahydrofolate reductase"

/codon_start=1

/product="At2g44160/F6E13.29"

/protein_id="AK91450.1"

/db_xref="gi:15215810"

/translation="MKYIDKIQSLADGCKIAFSEFFPKPTGDCDNLFEKRDRAVATGPFCDITWAGGSDADITLISRKNVVCVSMHMLCTINMPKPEIDHALETITRSR

5' UTR

CDS

| 3' UTR | | 525 a | 368 c | 460 g | 564 t |
|---|-----------|---|-------|-------|-------|
| BASE COUNT | | | | | |
| ORIGIN | | | | | |
| Alignment Scores: | | | | | |
| Prod. No.: | 1.35e-152 | Length: | | 1917 | |
| Score: | 1834.00 | Matches: | | 334 | |
| Percent Similarity: | 91.15% | Conservative: | | 37 | |
| Best Local Similarity: | 82.06% | Mismatch: | | 36 | |
| Query Match: | 83.86% | Indels: | | 0 | |
| DB: | 8 | Gaps: | | 0 | |
| US-09-720-451-6 (1-408) x AY050434 (1-1917) | | | | | |
| QY | 1 | AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn | 20 | | |
| Db | 599 | GCAGCATCTGATGTTCTCAGCTTTTCATGATCATGATATATCCCAAGTTTGGAAT | 658 | | |
| QY | 21 | AspCysArgGlnIleGlyIlePheTyrProGlyIleMetProIleAsnAsn | 40 | | |
| Db | 659 | GATTTCGGCAATTGGATTGATTGTCCTCATTTGCTCGGAATATAGCCTATTATATAC | 718 | | |
| QY | 41 | TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMet | 60 | | |
| Db | 719 | TACAGAGCTTTTGGTGGTACTGGTTTCGACAGACTAAGATACCAAGTTGGAGGATG | 778 | | |
| QY | 61 | AlaAlaLeuGlnIleProIleLysAspAsnGlnGlnIleValLysAlaTyrGlyIleHisLeu | 80 | | |
| Db | 779 | GCAGCCTTGAGCCATCAAGATTAACGAGAAGAGTGGAAAGCCATGGTATTCACCTT | 838 | | |
| QY | 81 | GlyThrGlnMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr | 100 | | |
| Db | 839 | GGAACACAGATGTGTAAAGATGTGGCTCATGGAGTCAAGTCTTCTCATCTTCACCA | 898 | | |
| QY | 101 | LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnGlnSer | 120 | | |
| Db | 899 | TTGACATGAGAAATCTGCTCTTGCAATATGATGATCTTGATGATGATGATGATGATC | 958 | | |
| QY | 121 | LysValSerArgSerLeuProTrrParGarProAlaAsnValPheValGlyValLysGlnAsp | 140 | | |
| Db | 959 | AAAAATTTCTGTTCTTACCCTGGAGAGCCCTGCAAAATGTTTCCGTACTAAGGAAGAT | 1018 | | |
| QY | 141 | ValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGlyTyr | 160 | | |
| Db | 1019 | GTGGGCCCATTTTCTTGGGCAACCGTCCAAAGACTCATTTCTTACAAACAGAGGGCTG | 1078 | | |
| QY | 161 | AspGlnTyrProHisGlyArgTrrPglLysAspSerCysAsnProSerTyrGlyAlaLeuSer | 180 | | |
| Db | 1079 | GAAATTTTCCCAACAGCCGGGGGTGATTACACCAAGTCTTATATGGTGCACTCTCG | 1138 | | |
| QY | 181 | AspTyrGlnPheMetArgProArgAlaAspLysLysLeuValGlnGlnTrrPalaVal | 200 | | |
| Db | 1139 | GATCATGAGTTTCCAGCTCCGGAGAGAGCTGCACAGAAGCTTCAACAAGAATGGGTGTC | 1198 | | |
| QY | 201 | ProLeuLysSerValGlnAspIleTyrGlnArgPheArgLeuTyrCysLeuGlyLysLeu | 220 | | |
| Db | 1199 | CCACTGAAAAGTGTAAAGTATTCACGGAATTTCAAGAGAGCTCTTGGAAACCTT | 1258 | | |
| QY | 221 | ArgSerAsnProTrrPserGlnLeuAspGlyLeuGlnProGlnTrrLysIleIleAsnGln | 240 | | |
| Db | 1259 | AAAAGCAGTCCCTGCTGTGAATAGTGGATCCAGCCAGACAGAAGATTAATAAACGAG | 1318 | | |
| QY | 241 | GlnLeuGlnLysIleAsnThrLysGlyPheLeuTrrIleAsnSerGlnProAlaValAsn | 260 | | |

Db 1319 CAACGTATAAAGTCACTCCAAAGGCTTCTTGACCATCATACCAACCATCATGATCAAC 1378
QY 261 GYGLULYSSERASPERPROTHRALGTYRPGYGLYPROGGLYTYRVALTYRGIN 280
Db 1379 GCCGAGAGATGATGCCCACTGCTGATGGGAGAGCTCTGTGATATGATACCA 1438
QY 281 LYSALATYRVALGLUPHETHECHESERYSGULYSLUENASPRALALEUVALASPLYSCYS 300
Db 1439 AAGCTTATCTAGAGTTCTTCTGCTCAAGAGAAATAGATGACGCGGAGAAATCC 1498
QY 301 LYSASPARGTHTSERLEUTHRTYMETALAYALANYSAPGLYSERTPLYSSERA 320
Db 1499 AAGCTTGCCATCGATCTTACATGCGCTGGAACAAGAGAACAGTGGATCGAAC 1558
QY 321 VALGVLGHTHRASPVALASNALAYALHTHRPGYALPHEPROALALYSGIULLE 340
Db 1559 ACTGCCCAACGCTGATGAAATCTTAACCTGGGAGTTTCCCGCTAGGAAATCAT 1618
QY 341 GLNPROTHRTLEVALASPROVALSERPHEASNAVALTRPYASPGUALAHPHEGIULLE 360
Db 1619 CAACCAACCATGTGCTGATCCCGCTTCAACGCTGGAAGATGAGCATTTGAGACT 1678
QY 361 TRSSEARGLYTRPALASERLEUTRYPROGLUASPGUALASERAGLYSLAUVALGU 380
Db 1679 TGGTCAAGAGAGCTGGCTACTGTACCCAGAACCCCTTCCAGAACTTGCTCGAG 1738
QY 381 GLUVALGLYGLYSETHSPHELEUVALSERLEUVALASPAASPLYRILEASNGLYASP 400
Db 1739 GAGGTGAAGAACGCTACTGATTTGGTGAAGCTTGGTGAAGACATTCATCATGATGATGAC 1798
QY 401 LEUPHEALAYALPHEALASP 407
Db 1799 ATATCGCGGTGTTCTGAT 1819

RESULT 3
AF370515
LOCUS AF370515 2034 bp mRNA linear PIN 30-APR-2001
DEFINITION Arabidopsis thaliana putative methylglutamate reductase
(At2g44160; FEEL3.29) mRNA, complete cds.
ACCESSION AF370515
VERSION AF370515.1 GI:13877628
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2034)
Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M.,
Palm, C. J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,
Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K.,
Sakano, H., Sakurai, T., Satou, M., Seki, M., Shino, P., Yamada, K.,
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R. W.
Direct Submission
Submitted (17-Apr-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arabseq@stanford.edu

COMMENT
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFLP cDNAs (RFLP cDNA: "RIKEN
Arabidopsis Full-Length cDNA"): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RFLP cDNAs: Southwick, A.,
Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M., Palm, C. J.,
Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K.,
Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shino, P.,
Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Southwick, A. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W.
(SSP/Stanford) contributed equally to this work as PIs.
location/Qualifiers
1..2034
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="11"
/clone="RAF11-04-D03"
/note="This clone is in plinescript
ecotype: Columbia"
1..2034
/gene="At2g44160; FEEL3.29"
91..1875
/cds
/gene="At2g44160; FEEL3.29"
/cds_start=1
/product="putative methylglutamate reductase"
/protein_10="AAK4382.1"
/db_xref="gi:13877629"
/translation="MKVLDKIQSLADEGKTAFSEFFPEKTEDGVNLFERRDMRAY
GPTFCDITWAGSGSTADLIDLSRQNVVCSIMHILCTMPEKIDHALETIRSN
GIQNYALRGDPHHGODKFOVEGDFDALDVNHRKYGDYFGITVAGYPEAHAPY
IGENGLASNEAYQSLLEYLKKIDAGADLITVQLFYDIDFLKFNOCROIGISCPY
PGIMPINNYRGELEMTGECKTIPVEMALPEIDKNEPAVAGVHIGTEPKMLA
HGVSILHYTLNMEKRSALATLMLGVIDSKISRLPMPRPANVRREEDRPTEFMAN
RPKYSISRTKMEDEFPQGRGDSRSAGALSIDHOPSPRADLQOEPVPLKSYE
DIQERFELGGLNLSKSPWSELDGLDPEIRINBOIKVNSKGLITTSQSVNARS
DSPVWGQGVGVYQKAYLEFCSKEKIDAVVEKKAIPSTTHMAVKGQWVSNRA
QADNAATWGVFPAKKEIIQPTIVDPSFVWMDAEAFETWSRMANLYEADPSRNLL
EVKNSYLVSVENDYINGDIFAVFDL"

BASE COUNT 568 a 386 c 483 g 597 t
ORIGIN

Alignment Scores:
Prod. No.: 1.46e-152 Length: 2034
Score: 1834.00 Matches: 324
Percent Similarity: 91.15% Conservative: 37
Best Local Similarity: 82.06% Mismatches: 36
Query Match: 83.86% Indels: 0
DB: 8 Gaps: 0

US-09-720-451-6 (1-408) x AF370515 (1-2034)

QY 1 AALARGVALILEVALTHCHLQENUPHETRYASPHRASPILEPHELEUNYSPHEVALAS 20
Db 649 GCTGATCTGATTTCTTACAGCTTTCTATCATCTGATATATCTTCAAGTTTGGAAT 708
QY 21 ASPCYARGGLNILEGLYLTHERCHSPROILEVALPROGLYILEMETPROILEASNAS 40
Db 709 GATTGCGGCAAAATGGGATAGTGTCCCATGCTTCTGGAATTAATGCTAATAATAC 768
QY 41 TYRVLGVLPHETLARGMETHCHLYPHECHYSLYSTRHLYSLIEPROMLAASPILEMET 60
Db 769 TACAGAGCTTTTGGCGTATGCTGTTTCCAGACATAAGATACCACTGAGTGAGATG 828
QY 61 ALALALEUGLUPROILELYSASPASNGIULNALAYALLYSALARYGLYLEHISLEU 80
Db 829 GCTGCCCTTGAGCTATCAAGATTAACGAAACCTGTGAAGCGTAATGATATTCACCT 888
QY 81 GLYTHGLUMECYSLYSLEYLLEUALHISGLYLYLESTRHLEUNHISLEUTRYTHR 100
Db 889 GGAACAGAGATGTGTAAGAAAGATCTGGCTCATGAGATCAAGTCTTCACTACACACA 948
QY 101 LEUASMETGLULYSERALEUALALEUWETASLEUGLYLEULIEGLUGLUSER 120
Db 949 TTGACATGAGAAATCTGCTCTTCATTAATGATGATGATGATGATGATGATGATG 1008
QY 121 LYSVALSERASERLEUPROTARYGATPROLAASNAVALPHEARVALYSGIULASP 140
Db 1009 AAAATTTCTGCTTTTACCTCGGAGAGCCCTGCAAAATGTTTCCGCTACTACGAAGAT 1068
QY 141 VALARGPROILEPHETRPALAASNARGPROLYSSERTYRILESERAGTHRTILEGLYTR 160
|||||

Db 1069 GTGCGCCCATTTTCTGGGCAACCGTCCAAAGACTACATTCTAGAACCAAGGCTGG 1128
Qy 161 AspGlnThrProHisGlyThrProGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
Db 1129 GAAGATTTCACAAAGCCGGTGGGTGATTCACGAGCTCTTCAATGAGTGCACCTCG 1188
Qy 181 AspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTrrPalAval 200
Db 1189 GATCATCAGTTCTACGTCGCGAGCAGCTGACAAAGAGCTTCACAAAGAAATGGGTGTC 1248
Qy 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
Db 1249 CCATGTAAGAGTGTAAATATATATATGATGATGATGATGATGATGATGATGATGAT 1308
Qy 221 ArgSerAsnProTrrPserGluLeuAspGlyLeuGluProGluTrrLysIleLeuAsnGlu 240
Db 1309 AAAAGCAGTCCCTGCTGCTGATTAATGATGATGATGATGATGATGATGATGATGAT 1368
Qy 241 GlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
Db 1369 CAATGTAAGAGTGTAAATATATATATGATGATGATGATGATGATGATGATGATGAT 1428
Qy 261 GlyLysSerAspSerProThrValGlyTrrPglYglYProGlyGlyTrrValTrrGln 280
Db 1429 GCCGAGAGATCTGATTCCTCCCACTGTTGATGGGAGGTCCTGTTGATGATGATGAT 1488
Qy 281 LysAlaTrrValGlnPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
Db 1489 AAGCTTATCTAGAGTCTCTCTGCAAGGAGAAATATGATGATGATGATGATGATGATGAT 1548
Qy 301 LysAspArgThrSerLeuThrTrrMetAlaValAsnLysAspGlySerTrrLysSerAsn 320
Db 1549 AAAGCTTGGCATTGATTTTACATGCTGTGAGAACAGAGGAGAACAGTGGATGATGAT 1608
Qy 321 ValGlnThrAspValAsnAlaValThrTrrPglYvalPheProAlaLysGluIleLeu 340
Db 1609 ACTGCCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1668
Qy 341 GlnProThrIleValAspProValSerPheAsnValTrrPglYAspGluAlaPheGluLe 360
Db 1669 CAACCAACCAATGCTGATCCGCTGATCTGATGATGATGATGATGATGATGATGATGAT 1728
Qy 361 TrrPserArgLysTrrPalAserLeuTrrProGluLysAspGluAlaSerArgLysLeuValGlu 380
Db 1729 TGGTCAGAGAGTGGGCTGATCTGTACCCAGAGCCGACCTTCAGAAAGCTGCTGAG 1788
Qy 381 GlnValGlyLysSerPheLeuValSerLeuValAspAsnAspTrrIleAsnGlyLysP 400
Db 1789 GAGGTGAAGAACAGTACTATTTTGGTAACCTTGTTGAGAACAGATTACATCAATGCTGAC 1848
Qy 401 LeuPheAlaValPheAlaAsp 407
Db 1849 ATATTTCGGGTGTTGCTGAT 1869

RESULT 4
ATH245414 2037 bp mRNA linear PLN 01-SEP-1999
LOCUS Arabidopsis thaliana mRNA for methylenetetrahydrofolate reductase.
ACCESSION AJ245414
VERSION AJ245414.1 GI:5823582
KEYWORDS 5,10-methylenetetrahydrofolate reductase; methylenetetrahydrofolate reductase; MTHFR1 gene.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2037)
AUTHORS Ravelle, S., Rebelle, F. and Douce, R.
TITLE Folate metabolism in higher plants: cloning of a cDNA for 5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2037)

AUTHORS Ravelle, S.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1999) Ravelle S., DBMS / PCV, URA-CNRS 576, CEA-Grenoble, 17, rue des Martyrs, 38054 Grenoble Cedex 9, FRANCE
FEATURES
source 1..2037
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
68..1852
/gene="MTHFR1"
68..1852
/gene="MTHFR1"
/EC_number="1.5.1.20"
/function="reduction of methylenetetrahydrofolate to methylenetetrahydrofolate"
/codon_start=1
/product="methylenetetrahydrofolate reductase"
/protein_id="CA53783.1"
/db_xref="GI:5823583"
/translation="MKVLDKIQSLADEKTAFFSFPPEPKTEGVDNLFERMDRVAY
GPTFCIDITWAGGSGTADLTLDIASRQSVVCSNMHLCTNMEVEKIDHALETIRN
GIONVALRGDPHGDGKFOVEGDFCDLADLVNIRSKYDYGIVAGYEPANPY
IGENGLASNEAYOSDLEYLKKIKDAGADLITQLPYDDIDFLKFNDRQIGSCPV
PGIMPINNYRGLRMGPFCKTKIPVEVMALEIRIDNEAVKAVGHILGTEMKKMA
HGKSLHLTYLNNKRSALILNMLGMDSEKISRELPPRRANPTREKEDYRPLKSYE
RKSYLSRTGMEIDFPQGRGDSRSASTGALSDRQFSRRARDDKLDQERYVPLKSYE
DIQEKELCLGMLKSPWSELDGLOPEIRIINEGLIKVNSKGLTINSQVNAERS
DSPVWGSGPVGYVYKAYLEFCSKEKLDAYEVCALPSTYNAVWKGQWNTVY
QADVNATWQVFPFAKEITDPTVDPASFSPWDEAFETWSSMNLYPEADPSHMLE
EVKSYLYLVEINDYINDIDFAVEFDL"

BASE COUNT 583 a 385 c 480 g 589 t
ORIGIN

Alignment Scores:
Pred. No.: 2,686-152 Length: 2037
Score: 1831.00 Matches: 334
Percent Similarity: 91.15% Conservative: 37
Best Local Similarity: 82.06% Mismatches: 36
Query Match: 83.72% Indels: 0
DB: Gaps: 0

US-09-720-451-6 (1-408) x ATH245414 (1-2037)

Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
Db 626 GGTGATCTGATGTTTACACACCTTTCTATGATATGATATATATCTCAAGTTGGAA 685
Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
Db 686 GATTGTGGCAAAATGGGATTAGTTCCTGATGCTGGAATTTATGCTATTAATAC 745
Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaIleMet 60
Db 746 TACAGAGGCTTCTTCGATACCTGTTTCGCAAGCTAAGTACAGTACAGTGGATG 805
Qy 61 AlaAlaLeuGluProIleLysAspAsnGluAlaValLysAlaTrrGlyIleHisLeu 80
Db 806 GCTGCTTGGAGCCCTCATACAGATACGAAAGAACTGTGAACCTATGATATTCACCT 865
Qy 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTrrThr 100
Db 866 GGAACAGAGATGTGAAAAGATGTGGCTCATGATGATCAAGTCTCTCATCTCACACA 925
Qy 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120
Db 926 TTGAACATGGAATAATTCGCTTCGCAATATGATGATGATGATGATGATGATGATGAT 985
Qy 121 LysValSerArgSerLeuProTrrPargArgProAlaAsnValPheArgValLysGluAsp 140
Db 986 AAAATTTCTGCTTCTTACCTCGAGAGCCCTGCAATGTTTCCGTACTAAGGAAGT 1045
Qy 141 ValArgProIlePheTrrPalAsnArgProLysSerTrrIleSerArgThrIleGlyTrrP 160

Db 1046 GTGCGCCCATTTCTGGCAAAACGCTCAACAGCTACATTCTAGAACCAAGCGCTGG 1105
 QY 161 ASpgLtrYrProHslGlyrTrpGlyAspSerCysAsnProSerYrGlyAlaLeuSer 180
 Db 1106 GAAAGACTTCCACACAGGCGCGTGGGGTGATTCACCCAGCTTCAATATGGGCACTCTGG 1165
 QY 181 ASprYrGlnPheMetarProArGAlaArGAspLysLysLeuValGluGluTrpAlaVal 200
 Db 1166 GATCATCAGTTCTCACGCTCCGACAGCTGACACAGAGCTTCAACAGATGGGTGTC 1225
 QY 201 ProLeuYrSerValGluAspLlyrYrGluArGpHeArGLeuYrYrCysLsengLysLeu 220
 Db 1226 CCACTGAAAAGCTTCAACATATTCAGAGAAATTCAGAGAGCTCTGCTTGGAAACCTT 1285
 QY 221 ArgSerAsnProTrpSerLeuAspGlyLeuGlnProGluTrpLysLleIleAsnGlu 240
 Db 1286 AAAAGAGCTCCCTGGTCTAAATTAATGACCCAGCAGAGAGAGATATATTAACGAG 1345
 QY 241 GlnLeuGluLysLleAsnThrLysGlyPheLeuThrLleAsnSerGlnProAlaValAsn 260
 Db 1346 CAAGCTGATAAAGTCAACCTCAAGAGCTTTCAGACATCAATACCAACCATGCTAC 1405
 QY 261 GlyGluLysSerAspSerProThrValGlyTrpGlyLysProGlyGlyYrValYrGln 280
 Db 1406 GCGAGAGATGATTCGCCCAACTGTGGATGGGAGGCTCTGTGATATGATACCA 1465
 QY 281 LysAlaTrpValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
 Db 1466 AAGGCTTATCTAGAGTTCTTCTGCTCAAGAGAAATTAATGACAGTGGTGAGAAATGC 1525
 QY 301 LysAspArGThrSerLeuThrYrMetAlaValAsnLysAspGlySerTrpLysSerAsn 320
 Db 1526 AAGGCTTGGCCATCGATTCATTCATGCTGAGCAAGAGAGAAAGAGGATGCTATCGAAC 1585
 QY 321 ValGlyLThrAspValAsnAlaValThrTrpGlyValPheProAlaLysGluIleIle 340
 Db 1586 ACATGTCAGAGCTGATGATGATGCTGATGCTGATGCTGCTTCCCGCTAAGAAATCAT 1645
 QY 341 GlnProThrLleValAspProValSerPheAsnValTrpLysAspGluAlaPheGluIle 360
 Db 1646 CAACCAACCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1705
 QY 361 TrpSerArGlyTrpAlaSerLeuYrProGluAspGluAlaSerArGlyLysLeuValGlu 380
 Db 1706 TGGTCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1765
 QY 381 GluValGlyGlySerHisPheLeuValSerLeuValAspAsnAspYrLleAsnGlyAsp 400
 Db 1766 GAGGTGAGAGACAGCTACTATTTGGTGAAGCTTGGTGAAGAGATCATCATCATGATG 1825
 QY 401 LeuPheAlaValPheAlaAsp 407
 Db 1826 ATATTCGCGCTCTTCTCTCAT 1846

RESULT 5
 AY122922 1810 bp mRNA linear PLN 03-JUL-2002
 LOCUS Arabidopsis thaliana putative methylenetetrahydrofolate reductase
 DEFINITION MTHFR1 (At3g59970) mRNA, complete cds.
 ACCESSION AY122922
 VERSION AY122922.1 GI:21689666
 KEYWORDS FLI_cDNA.
 SOURCE chile cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Rosidae; eudicots II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1810)
 Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,
 Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yu, G.,
 Bower, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,

Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.,
 Arabidopsis Open Reading Frame (ORF) Clones
 Unpublished
 2 (bases 1 to 1810)
 Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,
 Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yu, G.,
 Bower, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.,
 Direct Submission
 Submitted (17-JUN-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

COMMENT
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the pUNT (ORF) clones using the RAFL cDNAs: Yamada, K.,
 Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M.,
 Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yu, G., Bower, L.,
 Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B.,
 Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A.,
 Davis, R.W., Ecker, J.R. and Theologis, A.

FEATURES
 source
 location/Qualifiers
 1..1810
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="U18407"
 /note="This clone is in pUNT 51.
 ecotype: Columbia"
 1..1810
 /gene="At3g59970"
 1..1779
 /gene="At3g59970"
 /codon_start=1
 /evidence=experimental
 /product="putative methylenetetrahydrofolate reductase
 MTHFR1"
 /protein_id="AA067455.1"
 /db_xref="GI:21689667"
 /translation="MKVYDKTSVTEGCGTAAFSPEEPPEKTEGCVENLPERMDRLVSY
 GPRFDITMGAGSTADITLETASRMONVCEETMMLTETNNPIKIDHALETIRSN
 GIONVYALGDPFHGQDKFVVEGSRFADLVNHRISKTDGFGTGVGPAPHDV
 IADGATPEESQSDLAFLKRVAGADLVLTQDLEFLEKVDGCKQIGNCIV
 PGIMPISNKGFLRMAGFCRTKIPAEFLALEPIKIDENDEVRKAGVGHFAETMCKKLLA
 HGITSLHLYTLNVDSKSAIGILMNLGIDESKISLPRPAPAVPFTKEDVPIEWAN
 RKESYISRTKGWNPDPGRGDSHSAVSTSDYQFAPRPGKPKIOOEWVYPLKIS
 DVOEKPEKELICNLSKSPSELDIGQETKILNOLKATINSNELFINSNVNAKS
 DPAIGMGSGPGCYVYQKAYLEFRFSKDKLTIVKSKAFPSITTYMAVKNSEWVNTG
 ESDVNAVTVGVPFAKEVLPFTVDPASFKWCKWDEAFIMSRWANLIPEDDPSKILLE
 EVKNSYLVSLVDNNYINGDIFSVFA"
 1780..1810
 /gene="At3g59970"

BASE COUNT 508 a 372 c 427 g 503 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,37e-147 Length: 1810
 Score: 1777.00 Matches: 319
 Percent Similarity: 89.41% Conservative: 44
 Best local Similarity: 78.57% Mismatches: 43
 Query Match: 81.25% Indels: 0
 DB: 8 Gaps: 0

| | | | |
|------------------------|--------|---------------|----|
| Percent Similarity: | 89.41% | Conservative: | 44 |
| Best Local Similarity: | 78.57% | Mismatches: | 43 |
| Query Match: | 81.25% | Indels: | 0 |
| DB: | 8 | Gaps: | 0 |

| | | | |
|----|------|---|------|
| Qy | 1 | AlAaRgVallEvaLthGlnLeuPheArgAspTrpAspIlePheLeuPheValAsn | 20 |
| Db | 621 | GCAGATTTCATTGGACACAGCTTTCTGATACTGATATATTCCTCAAGTTTGAGAT | 680 |
| Qy | 21 | AspCyAsArgGlnIleGlyIleThrCysProIleValProGlyIleWetProIleAsnAsn | 40 |
| Db | 681 | GATTTCGGCCAAATCGGATTAATTTGCCATTGTTCCGTGAATTAAGCTATATTCAC | 740 |
| Qy | 41 | TyrTyrGlyPheIleArgMetThrGlyPheCysTyrTrpIleProAlaAspIleMet | 60 |
| Db | 741 | TACAAAGGGTCTTCGATATGGCGGTTCCTAAGACCAAGATACCCGTGAGCACT | 800 |
| Qy | 61 | AlAlAlaLeuGlnProIleLysAspAsnGlnGluAlaValLysAlaTyrGlyIleHisLeu | 80 |
| Db | 801 | GCAGCTTGAACCCATTAAAGGATTAATGACAGAGCGGTGTTAAAGCCCTGTGAATTCAC | 860 |
| Qy | 81 | GlyThrGluMetCysLysIleLeuAlaHisGlyIleTyrThrLeuHisLeuTyrTrp | 100 |
| Db | 861 | GCACAGCAATGTGCAAAAAAGATTTTGGCCATGAGAACACTTCCTCATCTCTACACA | 920 |
| Qy | 101 | LeuAsnMetGlnLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnGlnSer | 120 |
| Db | 921 | TTTAAGCTGGACAAATACGCTATTGGATTAATTAAGAACCTGTGCTGATGTATGAGTCA | 980 |
| Qy | 121 | LysValSerArgSerLeuProTrpArgArgProAlaAsnValPheArgValLysGluAsp | 140 |
| Db | 981 | AAATATTTCGTTCTCTACCTTGAGACGCGCTGCAAAATGTTTTCCGTACTAAGGAAGAT | 1040 |
| Qy | 141 | ValArgProIlePheTrpAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTrp | 160 |
| Db | 1041 | GTTTCGCCAATTTTGTGGCAAAACCGTCAAAAGACTTCATATTTGAACAAAGGGCGTGG | 1100 |
| Qy | 161 | AspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeuSer | 180 |
| Db | 1101 | AATGACTTCCACATGACGAGCGTGGGGGATTCACACAGTCGACATACACTGCTTTCG | 1160 |
| Qy | 181 | AspTyrGlnPheMetArgProAlaArgAlaArgAspLysLysLeuValGlnGlnTrpAlaVal | 200 |
| Db | 1161 | GATTATTCAGTTTGGCGGCCCAAAAGACGCGACAAAGAGCTTCCAGCAAGATGGCGTGC | 1220 |
| Qy | 201 | ProLeuLysSerValGluAspIleTyrGlnArgPheArgLeuTyrCysLeuGlnLysLeu | 220 |
| Db | 1221 | CCACTGAAGACATTGAAGATGTTCAAGAGAAATTCAAAGAGCCTTCATTTGAAACTTA | 1280 |
| Qy | 221 | ArgSerAspProTrpSerGlnLeuAspGlnProGlnProGlnTrpLysIleIleAsnGln | 240 |
| Db | 1281 | AAAGAGACGCCATGGTCTGGAATTAAGATGACTCCACCGACAGACAAAGATCATAAATGAG | 1340 |
| Qy | 241 | GlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn | 260 |
| Db | 1341 | CACACTCGGAAAAATCAATCCACAGCGTTCTCGACCATTAATAGCCAAACATCAGTCAAT | 1400 |
| Qy | 261 | GlyGlnLysSerAspSerProThrLysGlyTrpGlyLysProGlyLysIlyTrpValTyrGln | 280 |
| Db | 1401 | GCAGCCAAATCCGATTTCTCACTGACTATGATGGGTGGTCTCGTGGTTACGTTACCGAG | 1460 |
| Qy | 281 | LysAlaTyrValGlnPhePheCysSerLysGlnLysLeuAspAlaLeuValAspLysCys | 300 |
| Db | 1461 | AAAGCTTATCTGGAGTCTTCTCTGTTCAAGAGATTAAGTTAGACACACTTGGGAGAAATCC | 1520 |
| Qy | 301 | LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLysSerAsn | 320 |
| Db | 1521 | AAAGCTTTCCTCTTATGACTTACATGCGCGTGAAACAAATCAGAGAAATTTGGGATTCAAAC | 1580 |
| Qy | 321 | ValGlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLysGlnIleIle | 340 |

| | | | |
|------------|--|--|------|
| Dp | 1581 | ACCGGTAGTCCGATGTGAATCACTTACTTTGGGAGTGTTCCACCTAAGGAGGTATT | 1640 |
| Oy | 341 | GlnProThrIleValAspProValSerPheAsnValITrpyAspsptLnaIaPheGluLe | 360 |
| Dd | 1641 | CACCGCAGCATGTGGATGCCACCAGTTTCAAAGCTTGGAAAACAACACGTTTGAGT | 1700 |
| Oy | 361 | TpserArgIlyTrpaIaserLeutyRyrProGluAsptLnaIaseraArgysLeuValGu | 380 |
| Dd | 1701 | TGGTCAAGAAGCTGGGCTTAACCTTGTACCCAGAACATGCCTTCFMAAAGTGTCTCG | 1760 |
| Oy | 381 | GluValAlaLyserHisPheLeuValSerLeuValAspAsnAspTryIleAsnGlyASP | 400 |
| Dd | 1761 | GAGGTGAGAGACACGCTACTACTTGTGGTAGCTAGTGGACACATTCATCATGTCTGAT | 1820 |
| Oy | 401 | LeuPheaIaValPheaLa 406 ::: : | |
| Dd | 1821 | ATCTTCTCTCTTCTTCT 1838 | |
| RESU.F 7 | | | |
| LOCUS | AY070034 | 2002 bp mRNA linear PLN 26-Apr-2002 | |
| DEFINITION | Arabidopsis thaliana putative methyl-methyltetrahydrofolate reductase | | |
| ACCESSION | MHFR1 (At3g59970) mRNA, complete cds. | | |
| VERSION | AY070034 | | |
| KEYWORDS | AY070034.1 GI:17979046 | | |
| SOURCE | FLI CDNA. | | |
| ORGANISM | Arabidopsis thaliana. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi. | | |
| REFERENCE | (bases 1 to 2002) | | |
| AUTHORS | Yamada,K., Liu,S.X., Sekano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Tortiumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Seaton,R.W., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. | | |
| TITLE | Arabidopsis Full Length cDNA Clones | | |
| JOURNAL | Unpublished | | |
| REFERENCE | 2 (bases 1 to 2002) | | |
| AUTHORS | Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,G.S., Quach,H.L., Tang,C.C., Tortiumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (12-Dec-2001) Plant Gene Expression Center, 800 Buchanan | | |
| COMMENT | Street, Albany, CA 94710, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN Arabidopsis Full-Length cDNA) : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. | | |
| FEATURES | The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J., Chang,C.H., Chang,E., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,G.S., Quach,H.L., Tang,C.C., Tortiumi,M., Wu,H.C., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shin,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A. | | |
| source | Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs. Location/Qualifiers 1..2002 | | |


```
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAF109-25-111 (R18407)"
/note="This clone is in a modified pBluescript vector
(FIC-1) as a BamHI/XhoI insert.
ecotype: Columbia"
1..2002
/gene="At3g59970"
5'UTR 1..82
/gene="At3g59970"
CDS 83..1861
/gene="At3g59970"
/codon_start=1
/evidence=experimental
/product="putative methyltetrahydrofolate reductase
MTHFR1"
/protein_id="AA144971.1"
/db_xref="GI:17979047"
/translation="MKVVDKIKSVTEGQIAFSEFPPEKTEDEVENLPERMDLVSY
GPTFCDITWAGAGSTADLTLEIASRMQVCEVEMMLCTNMPIDIEIDALEIRSN
GIONVALRGDPHGDQKVEYGGFACALDVNHRISKYDYGITVAGYPAHNPY
IEADGLATPESYOSDLAYLKKVADADLIYOLFYDIDILAKFNDRCROIGINCPY
PGIMPSNKKGEIRMAAGFCETKIPALFALEPIKNDENAVANGIRHATEMCKILA
HGITSILYTLVNDKSAIGILMNLGLIDESKISLSLPPRRAPAKFEKDEVRPFMAN
RPSYSISRTKGNNDPFGHMGDSHSAVSTLSDFOPAPRRGRDCKLOOEVRPLKSE
DNOEKREKICGNLSSPSRSDIGLOPRTKINDICKINSCELTINSOPVNAKS
DSPAIKMGPRGGYVOKATLEFRCCKDLIDLVKSKAFPSITTMVANKSNVSNIG
ESDVNAVTVGVPFAKEVIOPTIVDPASFKWMDKAFETWSHMANLYPEDPSKLE
EVKNSTYLVLSVLDNNTINDISFVA"
1862..2002
3'UTR /gene="At3g59970"
BASE COUNT 571 a 409 c 450 g 572 t
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-147 Length: 2002
Score: 1777.00 Matches: 319
Percent Similarity: 89.41% Conservative: 44
Best Local Similarity: 78.57% Mismatches: 43
Query Match: 81.25% Indels: 0
DB: 8 Gaps: 0
US-09-720-451-6 (1-408) x AY070034 (1-2002)
```

```
QY 1 AAlaargValIleValIthrgInleuphetYrAspThrAspIlePheleuLysPheValAsn 20
111 ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 641 GCAGATTGTGATTGACTCAGCTTTTCTATGATGATGATATATCTCCTCAAGTTTGTGAAT 700
QY 21 AspcysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn 40
111111111111111111111111111111111111111111111111111111111111111111
Db 701 GATTGTGCGCAATCGGATTAATTGTCCTATGTTCTCGAGATATAGCTATTTTCANAC 760
QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
111111111111111111111111111111111111111111111111111111111111111111
Db 761 TACAAGGGGCTTCTGCTATGCGCTGCTTCTGTAAAGCAACCAAGATACCCCTGAGCTCACT 820
QY 61 AlaAlaLeuGlnProIleLysAspAsnGlnGluAlaValIleValIleGlyIleHisLeu 80
111111111111111111111111111111111111111111111111111111111111111111
Db 821 GCTGCTTACAGCCTATTAAAGATATAGCAGGCTGTATTAAGCCTATGGAATTCACCTTT 880
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisIleuThr 100
111111111111111111111111111111111111111111111111111111111111111111
Db 881 GCAACAGAAATGTGCAAAAGATTTTGGCCATGAAATCACTCCCTCATCTCTACACA 940
QY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleLeuIleuSer 120
111111111111111111111111111111111111111111111111111111111111111111
Db 941 TTGAACGGGACAAATCACTATTTGGATTTATATACCTTGCTGTGATATAGTCA 1000
QY 121 LysValSerArgSerLeuProIlePargArgProAlaAsnValPheArgValIleGlyLys 140
111111111111111111111111111111111111111111111111111111111111111111
Db 1001 AAAATTCTGCTCTACCTTGAGACGCCCTGCAAAATGTTTTCGTACTAGAGAGANT 1060
```

```
QY 141 ValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTyr 160
111111111111111111111111111111111111111111111111111111111111111111
Db 1061 GTTCCGCCAATTTTCTGGCCAAACCGTCCAAAGAGCTACTATCTGAAACAAAGGCTCG 1120
QY 161 AspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
111111111111111111111111111111111111111111111111111111111111111111
Db 1121 AATGACTCCCAATCAGTACGCTTGGGTGATTCACACAGTGCACATACATACATACATCTTCG 1180
QY 181 AspTyrGlnPheMetArgProArgAlaArgAspLysIleuValGlnIleThrPheAlaVal 200
111111111111111111111111111111111111111111111111111111111111111111
Db 1181 GATATACAGTTTGGCCGCCAAAGAGAGACAGACAGAGAGTTCAGCAGAGAGAGGCTCGTC 1240
QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
111111111111111111111111111111111111111111111111111111111111111111
Db 1241 CCACGAAAGATGAGATGATTCAGAGAGAAATTCAAAGAGCTGTCATGGAACCTTA 1300
QY 221 ArgSerAsnProTyrPseGluLeuAspGlyLeuGlnProGluThrLysIleLeuAsnGlu 240
111111111111111111111111111111111111111111111111111111111111111111
Db 1301 AAAGAGGCCATGTGCTGATTTAGATGAGACTCCAGCCAGACAGACATCAATTAATGAG 1360
QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
111111111111111111111111111111111111111111111111111111111111111111
Db 1361 CAACCTCGAAATATCACTCCAAAGGTTTCTCCTGACATTAATGACCAACATCAGTCAAT 1420
QY 261 GlyGluLysSerAspSerProThrValGlyTyrGlyIleProGlyGlyTyrValTyrGln 280
111111111111111111111111111111111111111111111111111111111111111111
Db 1421 GCAGCCAAATCCGATTCCTCCAGCTATGAGAGGGGTGCTGCTGCTGCTACCTACAG 1480
QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
111111111111111111111111111111111111111111111111111111111111111111
Db 1481 AAAGCTTATCTGAGATCTTCTGTTCAAGAGATTAAGTTAACACACTTGGGAGAAATCC 1540
QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspLysSerTyrLysSerAsn 320
111111111111111111111111111111111111111111111111111111111111111111
Db 1541 AAAGCTTTCCTTCTATACCTACCTACGCTGACCAAAATCAGAGATTTGGATTCAAAC 1600
QY 321 ValGlyGlnThrAspValAlaValIleThrTyrGlyValPheProAlaLysGluIleIle 340
111111111111111111111111111111111111111111111111111111111111111111
Db 1601 ACCGGTGAAGCTCGAGTAAATGACGATTAAGTGGGAGCTGTCCACCTAAGAGGATTAAT 1660
QY 341 GlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAlaPheGluIle 360
111111111111111111111111111111111111111111111111111111111111111111
Db 1661 CAACCCAGCATGTGATCCAGCCAGCTTCAAGTCTGGAAGACAGACAGCGTTTGAGAT 1720
QY 361 TrpSerArgGlyTyrPalaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
111111111111111111111111111111111111111111111111111111111111111111
Db 1721 TGTCTCAAGAGCTGCGCTACCTGTACCCAGAGAGATGACCTCTAGAAAGTTGCTCGAG 1780
QY 381 GlnValGlyLysThrIlePheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
111111111111111111111111111111111111111111111111111111111111111111
Db 1781 GAGGTCAAGAACAGCTACTATTGTTAGCTTAAGCTTACGACAAACATTCATCAATGTGAT 1840
QY 401 LeuPheAlaValPheAla 406
111111111111111111111111111111111111111111111111111111111111111111
Db 1841 AATCTCTCTGCTTGTCT 1858
```

```
RESULT 8
LOCUS AF174486 2225 bp mRNA linear PLN 13-DEC-1999
DEFINITION Zea mays methyltetrahydrofolate reductase mRNA, complete cds.
ACCESSION AF174486
VERSION AF174486.1 GI:5802605
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2225)
REFERENCE
Rojas, S., Wang, H., McNeill, S.D., Raymond, R.K., Applin, D.R.,
Shachar-Hill, Y., Bohner, H.J., and Hanson, A.D.
Isolation, characterization, and functional expression of cDNAs
encoding NADH-dependent methyltetrahydrofolate reductase from
```

higher plants
J. Biol. Chem. 274 (51), 36089-36096 (1999)
MEDLINE 20062814
PUBMED 10593891
REFERENCE 2 (bases 1 to 2225)
AUTHORS Wang, H. and Boehrert, H.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1999) Department of Plant Sciences, University of Arizona, Forbes Building, Room 303, Tucson, AZ 85721, USA

FEATURES
source
1..2225
/organism="Zea mays"
/db_xref="taxon:4577"
77..1858
/EC_number="1.5.1.20"
/function="catalyzes the reduction of 5,10-methylenetetrahydrofolate (CH₂-THF) to 5-methyltetrahydrofolate (CH₃-THF)"
/note="NADH-dependent; MTHFR"
/codon_start=1
/product="methylene tetrahydrofolate reductase"
/protein_id="AAD51733.1"
/db_xref="GI:5802606"
/translation="MKVTEKILAEAGDRTAFSPKPEPTKEGVEENLEPERMDRAVH GPSFCDITMGAGSTADLTLEIANRMGNMCAVETMMHLCNTMNPVKIKIDIALETTKSN GTONVLAIRGDPKPPKQKRVGEGFSCACNADIVQIIRAKYGFQGTIVKAGPEAPDA IOGEGGATLEAYSNDAIYKRVDAKDLVYDFIDTDFLEFVNDCKQITCEPIY PGIMPINNYKGRLEKMTGCKRIPELITADLPDKNEAVRQYGVGHIDTECKKILA TGKIRLHLVLTMDKSAIGILMNIGLIEESKVSRLPWRPRTNFRVKEDVPIRMAN RKRSYLKTRLTGMDOYPHGRWGDNRNPSGALTPHOFTPRGRKRLQDEMAVPLKSE DISERTNFCCGKLTSSPWSLIDGLQPEYKIIDQVANIQKGFLLINSQPAVNGEKS DSPVTGNGPGGYVOKAYLEFFCARKLDQLIEKTKAPSLYLIYNKDGTFENIS PNAVNAVTVFEPGEKIIOPVTVDHVASMKDEAFETIRMGCKNEPRGDSRELE KVKQTYIVSLVDNDYVGDLPFAFKI"

BASE COUNT 576 a 489 c 573 g 587 t
ORIGIN

Alignment Scores:
Pred. No.: 3,73e-146 Length: 2225
Score: 1762.00 Matches: 324
Percent Similarity: 88.64% Conservative: 35
Best Local Similarity: 80.00% Mismatches: 46
Query Match: 80.57% Indels: 0
DB: 8 Gaps: 0

US-09-720-451-6 (1-408) x AF174486 (1-2225)

QY 1 AAlaArgValIleValThrcInleupheryrAspThrAspIlePheLeuLysPheValAsn 20
||| :::
Db 635 GCTGACCTTATGTGTACACAACTTTCTATGATACCGACATCTTCTCAAGTTTGTGAAT 694

QY 21 ASpCCyArGInIleGlyIleThrcysProIleValProGlyIleMetProIleAsn 40
||||| :::
Db 695 GACTGCCGCCAAATGTGATCTGTCGCCATGTTCTGCGCAATTAAGCCAAATAAATAC 754

QY 41 TyrLyGlyPheIleArgMetThrcIlyPheCysLysThrIlyIleProAlaAspIleMet 60
||||| :::
Db 755 TACAAGGTTCTCTCGCATGACGAGGTCTGCAGAAATAGATACCTTCAGATCACT 814

QY 61 AAlaIleuGlnProIleLysAspAsnGluAlaValLysAlaIlyrGlyIleHisLeu 80
||||| :::
Db 815 GCTGACCTGATCTATCAAGACAAATGAGGCTGTTAACATATGAAATCCACCTT 874

QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrIleHisLeuTyrThr 100
||||| :::
Db 875 GGAACGTGAGATGTGCAAGAAATCTGTGCTACTGCGCATTAAGACTTTCGACCTTTACACA 934

QY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlnSer 120
||||| :::
Db 935 CTAAACATGACAACTCTCTATACGAATTTGATGAATCTTGATTAATTGAGAGACTCC 994

QY 121 LysValSerArgSerLeuProTrrArgArgProAlaAsnValPheArgValLysGluAsp 140
||||| :::

Db 995 AAGGTTTCAGAGCCATTACCTTGAGCCAGCGACTAAATGTTTCCGTTAAAGAGAT 1054
QY 141 ValArgProIlePheThrPalAsnArgProLysSerTyrIleSerArgThrIleGlyTrr 160
||||| :::
Db 1055 GTTCCACCTTATCTTGGGGCAACAGACCAAGACCTATCTTAAAGACATTTAGGTGG 1114

QY 161 ASpGInTrrProHisGlyArgTrrGlyAspSerCysAsnProSerTyrGlyAlaLeuSer 180
||||| :::
Db 1115 GATCAGATACCCATGGAGCGGTGGGTGATCTCGGAACCCATCATATAGAGACTTACT 1174

QY 181 AspTrrGlnPheMetArgProAlaArgAspLysLysLeuValGluGluTrrAlaVal 200
||||| :::
Db 1175 GACACCAAGTTCACAGACCCAGCGCGGTGTAAGAACTCCAGAGCAATGGGCTGT 1234

QY 201 ProLeuLysSerValGluAspIleTrrGluArgPheArgLeuTrrCysLeuGlyLysLeu 220
||||| :::
Db 1235 CCATCGAATCTGTGGAGGACATTAGAGCCCTTACAACTTCTGTCAAGGGAATCTC 1294

QY 221 ArgSerAsnProTrrPserGluLeuAspGlyLeuGlnProGluTrrLysIleIleAsnGlu 240
||||| :::
Db 1295 ACAAGCAGCCCATGCTCAGATTTGACGGCTCTTCAACAGACAAAGATTATGATGAC 1354

QY 241 GlnLeuGlnLysIleAsnThrLysGlyPheLeuTrrIleAsnSerGlnProAlaValAsn 260
||||| :::
Db 1355 CAGTTGGTGAATTTATACCAAGAGGTTCTTACATTAACACCAACCTGCTGTAAT 1414

QY 261 GlyLysLysSerAspSerProThrValGlyTrrGlyGlyProGlyGlyTrrValTrrGln 280
||||| :::
Db 1415 GGAAAGAAATCCAGACCTCCCTACTGTGGTGGGTGGTCCGAGGCTTACGTTATCAG 1474

QY 281 LysAlaTrrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
||||| :::
Db 1475 AAGGCTTACTCGAATCTCTCTCGCAAGAGAAAGAGTGGACCAACTATGAAAGATC 1534

QY 301 LysAspArgThrSerLeuThrTrrMetAlaValAsnLysAspGlySerTrrLysSerAsn 320
||||| :::
Db 1535 AAAGCAATCCCTCTCTCACTTACATGCTGTGAACAAAGAGAGGAACCAATCTCCAAAT 1594

QY 321 ValGlyGlnThrAspAlaAsnAlaValThrTrrGlyValPheProAlaLysGluIleIle 340
||||| :::
Db 1595 ATTTCACCGAAGCCGCGGATGCTGTCAGCTGGGTGGTTTCCCGCAAGAGATTAATC 1654

QY 341 GlnProThrIleValAspProValSerPheAsnValTrrLysAspGluAlaPheGluIle 360
||||| :::
Db 1655 CAGCCTACGGGTGGGTGATCGCAAGTTCATGATGTTGATGGAAGAGCAAGCAATTTGAATC 1714

QY 361 TrrPserArgIlyTrrPalaSerLeuTrrProGluAspGluAlaSerArgLysLeuValGlu 380
||||| :::
Db 1715 TGGACTCGGGGGTGGGTTCATGTTCCCTGAGGGTGAATCATCGAGGAGCTACTAGAG 1774

QY 381 GlnValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTrrIleAsnGlyAsp 400
||||| :::
Db 1775 AAGGTTCAAGAACTTACTTACCTGTGTACCTGTGTGACAGACTTACGTCAGGGGAG 1834

QY 401 LeuPheAlaValPhe 405
||||| :::
Db 1835 CTGTTGCTGCTTC 1849

RESULT 9
AC004005 109741 bp DNA 11near PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F6E13 map C1C10F02,
DEFINITION C1C02E07, complete sequence.
ACCESSION AC004005
VERSION AC004005.3 GI:20197067
KEYWORDS HMG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 109741)
AUTHORS Rounsley,S.D., Kaul,S., Lin,X., Ketchum,K.A., Crosby,M.L.,

Brandon, R. C., Sykes, S. M., Mason, T. M., Kerlavage, A. R., Adams, M. D., Somerville, C. R. and Venter, J. C.
Unpublished
2 (bases 1 to 109741)
Lin, X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 109741)
Town, C. D. and Kaul, S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@igf.org
On Apr 18, 2002 this sequence version replaced gi:6598402.
FEATURES
source
1..109741
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C10P02, C1C02E07"
/clone="F6E13"
complement(11..4149)
/note="overlap with BAC clone F18019 (AC002333:1..4149)."
complement(3191..3221)
/rpt_family="AT_rich"
3302..4750
/gene="At2g43880"
/note="F6E13.1"
join(<3302..3748,3846..4135,4219..4447,4532..>4750)
/gene="At2g43880"
join(3302..3748,3846..4135,4219..4447,4532..4750)
/gene="At2g43880"
/codon_start=1
/product="putative polygalacturonase"
/protein_id="AAC23426.1"
/db_xref="GI:3212875"
/translation="MANNISFPCALITFEFVNIIFLIKSHAMPFNVOYRGARGDRA
DARKSFLTMSLACGSRAAMVYVRGTYLVKNLFWGCKNIITFEKNGTLVAPAN
WDIGNSGYWILPAKYNRITSYGGTIDARAGYSGRCKHCPGARISFPCWNNVL
LISGLSFQSNHNVTVHSSNRIENRIKRAISGSPNIDGTHVOSSSGVTTISGCIAT
GDDCIALFSQSNRIWLEIERNCGPHGISIGSLGDYANEKVONVTVTSSVFTKTVNGV
RIKTVARSPRGFVNNVFNRLIMNNVENPVIIDQNYCPNGKCPROSSGVKLSGVTA
NIKGTSTPIAMKIDCSGSHCTGLRIDIKLYMRSSASYCRNAHGASGVMPRN
CM"

repeat_region
complement(5593..5613)
/rpt_family="AT_rich"
repeat_region
complement(6730..6753)
/rpt_family="AT_rich"
7192..8658
/gene="At2g43890"
/note="F6E13.2"
join(<7192..7632,7743..8032,8122..8350,8440..>8658)
/gene="At2g43890"
join(7192..7632,7743..8032,8122..8350,8440..8658)
/gene="At2g43890"
/codon_start=1
/product="putative polygalacturonase"
/protein_id="AAC23398.1"
/db_xref="GI:3212847"
/translation="MDNNKLIVAVLIMFESSFLMKSSSTASNYNVVSGAKPDGRDTS
TKAFLGAMQACRSAAVTVYRGSLPAPEFRGRCRSRTIFQYGTIVAPSDYRG
LGNNGYWILPVRNRIISITIGTLIDARAGSFMKRGSKCPYAGSMFPMANDVVS
GLTINSQTHLYVINCNNVYAKYKLVAPDOSPNIDGTHVOSSSGVTTISGCIAT
DCTISIGPGRNLTKLNCNGPHGISIGSLGDYANEKVONVTVTSSVFTKTVNGV
RIKTVARSPRGFVNNVFNRLIMNNVENPVIIDQNYCPNGKCPROSSGVKLSGVTA
NIKGTSTPIAMKIDCSGSHCTGLRIDIKLYMRSSASYCRNAHGASGVMPRN
CM"
gene
complement(9101..14090)
/gene="At2g43900"
/note="F6E13.3"
complement(join(<9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11384,11587..11671,

11768..11895,12004..12811,12910..13162,13259..>14090))
/gene="At2g43900"
complement(join(9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11384,11587..11671,
11768..11895,12004..12811,12910..13162,13259..14090))
/gene="At2g43900"
/codon_start=1
/product="putative inositol polyphosphate 5'-phosphatase"
/protein_id="AAC23399.1"
/db_xref="GI:3212848"

/translation="MDIINNHDENDDDEEALSMSSVPPPRKTHSHQRLATGQ
KGNHRQHSLLDIPKITEIVSCGISGSDSEDFPVATTTNSSEFTPGDTGSD
DYHLOPEIGEDFQPLPEFVSGGVMFVYRSPHSAKRPDLERPLPLETQDGR
FLNRIACETQLWAGSGEGRFENFDPAEPGCGSGRGVORGEDAPPOESASTPT
TCLMDNGNLWAGSGDKGRIRSMKMDYLDODDSDPFKEGLAQMHKGPVNSVMS
YGLMGSREGGVKIWTWESMEKSLRLEEKEMALLVRSIGDIDRAQVYNGTNI
SSSEYKCLADNVRSKYVMAKOLOTPELMGKREIKLIVNSSEGTENRMDPLGQOP
AAEDEMKAKTASTSKKEKPHGFLQSRNALMGADAVRRVATRGGAAYDARATEMY
LADGMITWGTCTGGLLIQWDGNGNRUDERHHQCAVLCCTREERTIGTVSGHIOI
DLBGNLTAGWVAHNNNAVIKMAADGYTFSLATHGGIRGMPVISPGLDIINSELAK
ERTYAQTDVRIILTGSWNVGQKASHDALSMVLSYADVGLIYVGLQEVMGAGFLA
MSAKESVGGNEGSTIGQYWIDTIGTLIDEKAVFERMGSRQLAGLLISLVRKRLTH
YGDIDVAAPCGGFRALGNKGVGRLRVPRDRIMCFINCHLAHLBAVNRNADPFHI
YKINSTRSRNNAHNAAPAGVSTGSHTTKSNMNAVMTEETKODLADADVVEGDDNY
RLRGTISYDEARDVFSQSFDMLEKQDLRAEMKAGVFQGMRAITTPPTTYKFEHR
PGLAGYDSEKERRIPAWCDRVIFRDRITSESSCSLDCPVASIMTYDCMDTESDH
KPVCKPFAVKLEHDEYRROEGRITIKTEKRYALINDLRVPEITVSNSTLVLDNQ
DTFVLRITNKCVKENAVERILICGEGSTVREDEPTELEKALIVYVGGCSTEVCHRV
KPDSSVEVSHHEEFHTELEFVDGIPQNMWCEPTRRKEALILVYVGGCSTEVCHRV
HYRRCESAKULRIDSNNSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNS
KSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNS
CSKQKSDSDDTNKSQKGDGSSSKSHKKNKGDDSSKSHKKNNDGSSKSHKSDG
DSSSKSHKSEGSLSLTRRTEILLQAHSTRRYGRKTVROL"

repeat_region
14103..14213
/rpt_family="GAGAA)n"
complement(14748..17292)
/gene="At2g43910"
/note="F6E13.4; supported by full length cDNA:
Ceres:24370"
complement(join(<14748..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..>17292))
/gene="At2g43910"
complement(join(14925..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..17218))
/gene="At2g43910"
/codon_start=1
/product="putative methyl chloride transferase"
/protein_id="AAC23400.1"
/db_xref="GI:3212849"
/translation="MAEEQNSDQSGNGNVLPPEVATFLHKTVEGGEKWEETI
TPMDGKATPLIVLVDPTSSLPGLRALVPGGGGHDVYAMASPRFVAGDISALSA
KANETYGSSPRAEFYSFYKEDVPTWRPELEFDLVFPCATPEPRAPMAKSMYEL
LKPGELITLWPTITDHYGPPYKVDVSTFEVYVLPIGFAVSEENPHALIPRKKE
KLGRWKKN"
repeat_region
complement(15108..15141)
/rpt_family="AT_rich"
repeat_region
complement(16570..16608)
/rpt_family="AT_rich"
repeat_region
complement(17642..17687)
/rpt_family="AT_rich"
repeat_region
complement(1774..17794)
/rpt_family="AT_rich"
repeat_region
complement(17926..17970)
/rpt_family="AT_rich"
repeat_region
complement(18152..18196)
/rpt_family="AT_rich"
repeat_region
complement(19342..21515)
/gene="At2g43920"
/note="F6E13.5"
complement(join(<19342..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,

/note="several differences to EMBL: X81973: Contains
ATP/GTP-binding site motif A (P-loop) AA169-176; Cysteine
synthase/cystathionine beta-synthase P-phosphate
attachment site AA143-161"
/codon_start=1
/product="cysteine synthase"
/protein_id="CA875795.1"
/db_xref="GI:7019670"
/translation="MYAMIASRPNREAKLASOILSTLLGNRCYSMAATSSALL
NPLSSSSSTLRPRPCSPSEISLSLSSSDSLAKKRSRFAQSEDEPSVCEAY
KRETPDGNATIDNWSQLGKTPMYLINSIANGVANIATLEIMEPCSVKDRIGYS
MYTDAQKGFISPGSVLVEPTSGNTGILAFIASSRGYRLITLTPASMSMERVLAK
AFCAELVLPDPAKMGAVQAKAEILIKNTPDAYMQQEDNPANPKIHETTPETLMD
TGKVDIFVADGIGTGTITGVGRFIEKKNPTQVIGVEPTESDILSGGPGPHKLOGI
GAGFIKNDOKIMDEVIAISSEAEITAKOLAKGKLWVGITSSGAAATAAKVAKRP
ENAGKLIATVNESNAVOYVFPSPGEXETLPTLPQSIREEVENQPEV"
/complement(5498..5626)
exon
/number=1
complement(5627..5839)
/number=1
complement(5840..5920)
/number=2
complement(5921..6041)
/number=2
complement(6042..6101)
/number=3
complement(6102..6252)
/number=3
complement(6253..6332)
/number=4
complement(6333..6431)
/number=4
complement(6432..6483)
/number=5
complement(6484..6610)
/number=5
complement(6611..6748)
/number=6
complement(6749..7010)
/number=6
complement(7011..7274)
/number=7
complement(7275..7404)
/number=7
complement(7405..7462)
/number=8
complement(7463..7548)
/number=8
complement(7549..7658)
/number=9
complement(7659..7818)
/number=9
complement(7819..8175)
/number=10
12111..18504
/gene="F24616.40"
complement(join(12111..12392,12490..12723,12815..13255,
13361..13588,13679..14482,14572..15081,15220..15462,
15581..15751,16213..17432,17543..17675,17797..18392,
18486..18504))
/note="contains ATP/GTP-binding site motif A (P-loop)
AA518-525; WW/rsp5/WMP domain signature and profile
AA518-540"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA875796.1"
/db_xref="GI:7019671"
/translation="MDLHPGGSKRKTSVVVTLDTGEEVYIASLLSKADQVIYIDPT
TGLIYNGKPGIDNFKSERELADYITNGSRGVSSVYARAIIIGYAVIGSGMLVAT
RLNPSIDPLPGGCVYVAESQWKIPLYNPQPGKGFTKIOBELTLDIDGKHYVAT
TRDITRPPSRMPLQSPDDEFVNNRWLSVPRKNTGLDEHCYIILIOGRAEVRPSSSCQ
LEGVIALMARSRRLHPGTRRYLAKGINSCTGNEVECDLWIKRNGQSIATFNSYIY
RGGTIPWGAELKMTAAEIVYADBDPKYGSTYEYQRLSKRYDTNLDAPVENOK

KKAFPIVCVNNLLSGSGKSCITIVQFPESAMNTRKSGGLPTRVHLIYDMHASYK
LKGRQITGEGIMWLKSPYATIGISEDDLPSSRNLKDCGEVYICIDIDGACLRSH
QNGVIRENCADSLDRTNMAASFPGLOVFEVRCRLGISLRTDLOYGNSVNNQGYNA
PLPGEWRKADAVGKSYIDHNKRTTWSPHCPDKPMKRLDMRFEFKSTILSPYS
ELADLFLOQGDIFHATLYTGSKAMHSIINFVPEBSGAKQFSAOKMKITLQRRYK
AMVSSROKOLEMFLGMRLEFKLPSIFVOPHLICPOAADIVELFIYSEPCYCOLL
TISHGADDLTCSPTVDYRTGHTIEDLKIYVNGASIPCAADCTNLYPLRPISSEMA
VTGGARLHEMDTSSLSLIDFEELBEQDLFTRVAVITTPAGAVRIPWTLQIEVL
GTSLPWGMFTCENTGRILAEARKPDEDSPPSSCDLNPFAATYSLOAEVSTVPQ
QKDPFNSNLIDLNGEDSSDPPEQPVVEIASGNDMLFDEAVVEYSGSDVPDQ
SVPONKRPKDSGAHLIYNLCISLGMVMVQNYFPVATMCCSDCMVSLAKLFEVE
AMKLEIRLRNLSIAERDRLSLIGIDPITINSXYDELTYIGRLCTANALAVNGO
ASLDDITIASIGLEKLENNVITRIBEGDGMGCYVRAVNSPVGSSSRG
ESSGVFLCPOCMKLENNVITRIBEGDGMGCYVRAVNSPVGSSSRG
KCCSSVFLQELIYDYRVNMSLRSGRVNAGREALEVEGNTNHLAVRQSPN
REDNPLKQILIGKEESIIEFPFSLKRVETADNSAFESLTLPLNLASSNAWKAAP
SADSEVAEIVLNTLSDVSVILIVSPGVDADAPOVQIASSDINKERTLNGKNDY
QSFTRSSPELSESGSGARPHIKFAKNVRCRITITLRLSGSSSVLDKNIN
LISDENPFAPIPRASFGATIEINDPCIHAKHILIVTWRDRTQSVESMSYRMD
RAPLRNPLIPIETERPMDIVLEIYLOASPLAGFRUDASATKPRYTHSPSSDQ
VDIWDPTSVIMEDRHVSFALITYIOVSYLQQYKVTIAYRIELDEARDGKULYDFPKQ
IQQRVFSKLDGVAATFDEPAEAVDLSRASPAAGLSIANRIKLYYADPEVGKMT
TSLSSV"

exon
/number=1
complement(12111..12392)
/number=1
complement(12393..12489)
/number=1
complement(12490..12723)
/number=2
complement(12724..12814)
/number=2
complement(12815..13255)
/number=3

intron
exon
intron
exon

Alignment Scores:
Pred. No.: 2,95e-114 Length: 100906
Score: 1424.00 Matches: 320
Percent Similarity: 47.51% Conservative: 42
Best Local Similarity: 41.99% Mismatches: 43
Query Match: 65.11% Indels: 360
DB: 8 Gaps: 8

US-09-720-451-6 (1-408) x ATF24G16 (1-100906)

QY 1 ALaarGvAlIleVAlThGInLeuPhenTyraSPHlePheLeuLysPheValAsn 20
||| :::
Db 84877 GCAGTTTGATTGATGACACACCTTTCTTATGATGATATCTCAAGTTTGTAAT 84936
QY 21 AspCysArGInIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
||||| :::
Db 84937 GATTGTGCGCAATCGGATTAATTTGCCATTTGCTTCTGGAATTAATTCATTTCCAAC 84996
QY 41 TyrlYsGlyPheIleArGmetThrGlyPheCysLysThrIysIle----- 55
||||| :::
Db 84997 TACAAGGGGTTCTGCGATAGGCTGTTCTGTAAAGCAAGGT-ATGTTGATCTCTT 85055
QY 55 ----- 55
Db 85056 TCATCTGTGTGATGATGATGCTTGTCTGTAGACACATGTCGCTTTGCTTAAAGGCT 85115
QY 55 ----- 55
Db 85116 TTAAGATGAGATTGAGTTTACTGCATCTATCTAGTTCCTTGGTGANATATATAGACC 85175
QY 55 ----- 55
Db 85176 TTGGTGATGATGCTTATGACTAATGCTGTCAGATATCATTTTCTTGTATTAATTC 85235
QY 56 -----ProAlaSPleIleMetAlaAlaLeuGluProIle 66
||||| :::
Db 85236 ACTCATGATCTTATTTATGATAGATACCGCGTGAGCTCACTGCTGCTTAAGCCTAAT 85295
QY 67 LysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMetCysLys 86

```

Db 85296 AAGGTAATATGACGAGCTGTTAAAGCTATGCAATTCCTTGCAACAGAAATGGCAA 85355
      |||
QY 87 LysIleuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsnMetGluLysSer 106
      |||
Db 85356 AAGATTTGGCCATGAAACACTCCCTCATCTCTACATGACATGACGCGCAAAATCA 85415
      |||
QY 107 AlaLeuAlaIleLeuMet----- 112
      |||
Db 85416 GCTATGGGATATATATGTTTCTCTCATTTCTCTTAATTTGGTTTCTGTGTAGTAA 85475
      |||
QY 113 -----As 113
      |||
Db 85476 TATTTACACATGCTATGTTGTATACCGGAGAACTAACAGTGTGCTCAATTCACAA 85535
      |||
QY 113 nLeuGlyLeuAlaLeuGluSerLysValSerArgSerLeuProTPrpArgProAlaAs 133
      |||
Db 85536 CCTTGTCTGTATGATGATGCAAAATTTCTGTTCTGTACCTGTGAGAGCGCCGCAAA 85595
      |||
QY 133 nValPheArgValLysGluAspValArgProIlePheTrp----- 146
      |||
Db 85596 TGTTCGCTACTAAGGAGAGATGTGCCCCAATTTCTGTAACTGATTTGTTTCTC 85655
      |||
QY 146 ----- 146
      |||
Db 85656 TTGCTCGAGATATATACAGATCTTAACTTAAGATATATTTGGCGGAGAAATTAATCT 85715
      |||
QY 147 -----AlaAsnArgProLysSerTyrIleSerArgThrI 158
      |||
Db 85716 GTTACTTCATTTTATCTTTGCGAGGCAACCGTCCAAAGAGCTACATATCTAGAACAA 85775
      |||
QY 158 LegIlyTrpAspGlnTyrProHisGlyArgTrpLysPserCysAsnProSerTyrGlyA 178
      |||
Db 85776 AAGGCTGAAGACTCTCCACATGAGAGCTGGGTAATTCACACAGTGCATACAGTA 85835
      |||
QY 178 IalLeuSerAspTyrGln----- 183
      |||
Db 85836 CACTTCGATTAACAGTACAGTACACACATATTTACAGTTTCGCATATCGACTTAA 85895
      |||
QY 183 ----- 183
      |||
Db 85896 TATTTTGAACCTGAAATTCGAAAGGCTGTTTTCAGTTTGCCCTTCGAAATTTGCTTAA 85955
      |||
QY 183 ----- 183
      |||
Db 85956 ATTGATATTGCATATTTGTTATTAGCCCTACATTAATCACTGTAGAAAGTGTGCTGT 86015
      |||
QY 183 ----- 183
      |||
Db 86016 TGCTTGCATCTCTTTCGCTGTTTTCATTCGCAATGAGTAGTGTATTAATGCGATGT 86075
      |||
QY 184 -----PheMetArgProArgIalArgAspLysLysLeuValGluGluTrpAlaValP 201
      |||
Db 86076 GTTGTATATTGTCGCGCCCAAAAGAGCGTGACAAAGAGCTTCGCAAGAAATGGTGTGTC 86135
      |||
QY 201 nLeuLysSerValGluAspIleLysTrpLysPheArgLeuTyr----- 215
      |||
Db 86136 CACTGAAAGCAATGAGATGTCAAGAGCT---AAGCTTACTTTACGACACTCGTA 86191
      |||
QY 215 ----- 215
      |||
Db 86192 CAATTTGTTTGCATCTTTCCTCATACTACCTGCACTTTTCATGTAAAAACTGTCA 86251
      |||
QY 216 -----CysLeuGlyLysLeuArgSerAsnProTPrpSerGluT 228
      |||
Db 86252 CAATTCAGAAATTCAAAGAGCTCTGCAATGTGAAACTTAAATAACAGACCGCATGTGAAAT 86311
      |||
QY 228 euAspGlyLeuGlnProGluThrLysIleIleAsnGluLeuGluLysIleAsnThrL 248
      |||
Db 86312 TAGATGAGACTCCAGCCAGAGCAAAAGATCAATTAATGTAGCAACTCGGAAAAATCAACTCA 86371
      |||
QY 248 y6GlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProT 268
      |||

```

```

Db 86372 ACGGTTTCCTGACCATTAATATAGCCACCATGATGATGACGAGCAAAATCGATTCCAG 86431
QY 268 hVal----- 269
      |||
Db 86432 CTATTGTGAAGAAAATCAAACTTGATGATGACAAAGCATACATTAACGTTCCAAAAA 86491
QY 269 ----- 269
      |||
Db 86492 AATCTTATTGACTGTGATTCACTGCTATTCTTAAAAACCTATTAAACCGCTGAG 86551
QY 269 ----- 269
      |||
Db 86552 ACGTTCTGTCTTACGTTATACATATGTTTAAATTCGATGGGAATATCTTGGCTT 86611
QY 270 -----GlyTrpGlyLysProGlyLysTyr 277
      |||
Db 86612 ATTAACAAAATGTATTAATTCGAATTGATCTTTCAGAGATGGGGTGTCTGTGTTTAC 86671
QY 278 ValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297
      |||
Db 86672 GTCTACCAAGAAAGCTTATCTGAGATCTTCTGTTCAAAGAGATTAAGTACACACTTGTG 86731
QY 298 AspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrp 317
      |||
Db 86732 GAGAAATCCAAAGCTTTTCTCTTATACCTACCTACATGCGCGGTGAACAAATCAGAGATTGG 86791
QY 318 LysSerAsnValGlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLys 337
      |||
Db 86792 GTATCAAAACCGGTGAGCTGCGATGTGATGACATTTCTTGGGAGAGTGTCCACGTTAA 86851
QY 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTrpLysAspGluAla 357
      |||
Db 86852 GAGGTATTCAACCCACCATCTGTGATCCAGCCAGTTTCAAAAGTCTGGAAAGCAGAGG 86911
QY 358 PheGluIleTrpSerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgLys 377
      |||
Db 86912 TTTGAGATTTGGTCAAGAGCTGGCTTAACCTTACCCAGAAAGTACCCCTTGTAAAG 86971
QY 378 LeuValGluGluValGly----- 383
      |||
Db 86972 TTGCTGAGAGAGGTAAA-GAATCTTCTCTTTTCTTCTCTTCTTTCAGTACAGAT 87030
QY 384 -----GlySerHis 386
      |||
Db 87031 TGCTTACCTTAATGATGATGTATGATCCATCGAATATGACGATGAGAAACACTAC 87090
QY 387 PheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAspLeuPheAlaValPheAla 406
      |||
Db 87091 TATTGTGAAGCTTGTAGTGACACAAATTAATCAATGATGTGATCTCTGTCTTGTCT 87150
      |||

```

RESULT 11
 AF441241 1785 bp mRNA linear syn 01-AUG-2002
 LOCUS AF441241
 DEFINITION Synthetic construct chimera1 mRNA, complete cds.
 ACCESSION AF441241
 VERSION AF441241.2 GI:22040812
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (bases 1 to 1785)
 Metabolic engineering in yeast demonstrates that
 methylecetrahydrofolate reductase reaction in vivo
 U. Biol. Chem. 277 (6), 4056-4061 (2002)
 2 (bases 1 to 1785)
 Roje,S., Raymond,R.K., Applling,D.R. and Hanson,A.D.
 Direct Submission
 Submitted (29-Oct-2001) Horticultural Sciences, University of

REFERENCE 3 Florida, Hull Road, Gainesville, FL 32611, USA
(bases 1 to 1785)
AUTHORS Roje,S., Raymond,R.K., Appling,D.R. and Hanson,A.D.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2002) Horticultural Sciences, University of
Florida, Hull Road, Gainesville, FL 32611, USA
REMARK Sequence update by submitter
COMMENT On Aug 1, 2002 this sequence version replaced gi:19550442.
FEATURES
source location/Qualifiers
1..1785
/organism="synthetic construct"
/db_xref="taxon:32630"
1..1785
/note="yeast-Arabidopsis hybrid methylenetetrahydrofolate
reductase"
/codon_start=1
/transl_table=11
/product="Chimeral"
/protein_id="AAL91367.2"
/db_xref="GI:22040813"
/translation="MKITEKLEQHORISGRPTYSFEYFVFKTTQGVONLYDMDHMYE
ASLPQIDITWAGGRLSHLSTDLVTAQSVLGLETCHMLTNTNPIIMIDALENA
YHSGGONLIALGDDPRDAENWTFEGGFOYADLIRKISYKHDFATGVAEGECH
PELPKDYKIDLELYEYKOKIDAGDPITTCOMFYVDNFTIMCSQVRAAGMDVPIIGIM
PTTYAALFRAQNGQISIPHPSRRIDPKDDDELVRDGTNLITYEKOKIDSGYV
SHDITMKLEKAPLMLERLNTLPTSERNAHPLAVLPKRSLNKRKNEEVRFLW
ANRPSTYSRTGMDPFHGWSDSHSAAYSTUSDQFAPRKGRDKLQOEENVPKLS
IEDVQEKELICIGNLKSSPWSLELDGAPETKLIINDELKINSNGFLITNSQPVAA
KSDSPAIGWGGGGYVYOKALFEFCRSKDKLIDLVESKAPFITTYAVNKSSENVSN
TGESDVNATWGVPAKKEVIQPTIVDPAKRVWMDAEFLIMSRMANLYPEDDPSRKL
LEEVKSYLVSLVDNNYINGDIFSVA"
BASE COUNT 506 a 415 c 424 g 440 t
ORIGIN
Alignment Scores:
Pred. No.: 2,92e-115 Length: 1785
Score: 1411.00 Matches: 258
Percent Similarity: 78.00% Conservative: 61
Best Local Similarity: 63.08% Mismatches: 84
Query Match: 64.52% Indels: 6
Gaps: 3
US-09-720-451-6 (1-408) x AF441241 (1-1785)
Qy 4 TLevalthrGlnleupheryrAsprhrAspIleleuLysPheValAsnAspCyArg 23
Db 556 ATCATCACTCAGATGTTTACGATGTTGATATTCATCACTCACTGCTTCCACAGTAGA 615
Qy 24 GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAspTyrGly 43
Db 616 GCTCGGGCATGCAGCGCCATTATCCCGGATCATGCCGATCAGCTACCTACGCGGCC 675
Qy 44 PheIleArgMetThrGlyPheCysTyrThrLysIleProAlaAspIleMetAlaIleu 63
Db 676 TTTCTGAGAAAGGCCCAATGCGCCAAATCTCCATCCCTCAACATTTCTCTCCGAGT 735
Qy 64 GluProIleLysAspAngIuAlaValLysAlaTyrGlyIleHisLeuGlyThrGlu 83
Db 736 GATCCTTCAAGCAGCATGCTGCTGCTGATTCGGAACGAACTTATATCGTGGAA 795
Qy 84 MetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyrThrLeuAsn 102
Db 796 ATGTGTCAAAAATGCTCGACAGTGTCTACGTTCTCACTGCAATCTACACCATGAAAC 855
Qy 103 MetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIle---GluGluSerLys 121
Db 856 TTGAAAAAAGCGCTCTCATGATTCGGAAGATTCACATTCACACGAGATTCAGAG 915
Qy 122 ValSerArgSer-----LeuProTTPArgArgProAlaAsnValPheArgVal 137
Db 916 TTCATATCACATCATTTGGCGGTGTCATGAGAAAATCTTGAATCCAAAGCGTAAA 975
Qy 138 LysGluLysValAlaArgProIlePheTrpAlaAsnArgProLysSerTyrIleSerArgThr 157

Db 976 AACGAGAACTCGACCTTATCTTGGCGAAACCGTCCAAAGACTCATATCTAGAAC 1035
Qy 158 TLeGlyTrpAspGlnProHisGlyArgTrpGlyLysSerCysAsnProSerTyrGly 177
Db 1036 AAGGCGTGAATGACTTCCACATGAGAGCTGGGTGATTCACAGATGAGATACAGT 1095
Qy 178 AlaIleSerAspTyrLysIlePheMetArgProAlaAlaArgAspLysLysLeuValGluGlu 197
Db 1096 ACACCTTGGATATATCAGTTTCGGCGCCAAAAGAGCTTCGAAAGAGCTTCGACAA 1155
Qy 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1156 TGGGTCTCCCATGAAAGCATTTGATGATGTTCCAGAGAAATTCAAAGAGCTCTCAT 1215
Qy 218 GlyLysLeuArgSerAsnProTrpSerGluLeuAspGlyLeuGlnProGluThrLysIle 237
Db 1216 GGAACCTTAAAGAGCGCCATGCTGATAGTAGTGAATGAGCTCCAGCCAGACAAAGATC 1275
Qy 238 TLeAsnGluLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnPro 257
Db 1276 ATAAATGAGCACTCGGAAAAATCAACATCCAGCGGTTCTGACCATTAATAGCAACCA 1335
Qy 258 AlaValAsnGlyLysSerAspSerProThrValGlyTrpGlyLysProGlyGlyTyr 277
Db 1336 TCATTCATACACCAAAATCCGATTCACAGCTATTGATGGGTGGTCTGCTGCTGTTAC 1395
Qy 278 ValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297
Db 1396 GTCTACACGAAACCTTATCTGAGCTCTCTGTTCAAGGATTAAGTTAGACACTGTG 1455
Qy 298 AspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrp 317
Db 1456 GAGAAATCCAAAGCTTCTTCTCTATACCTACATAGCGCGTGAAACAAATCAGAAATG 1515
Qy 318 LysSerAsnValGlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLys 337
Db 1516 GTATCAAAACACCGGTAGTCGATGATGATGATGATGATGATGATGATGATGATGAT 1575
Qy 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAla 357
Db 1576 GAGCTTATTCACCGACGATCGGATCCAGCATTTCAAAAGCTCTGGAAGACGAAGCG 1635
Qy 358 PheGluIleTrpSerArgGlyTyrPalaSerLeuTyrProGluAspGluAlaSerArgLys 377
Db 1636 TTTGAGATTTGGTCAAGCAAGCTGGGCTAATCTGTACCCAGAAATGACCTTACAGAAAG 1695
Qy 378 LeuValGluGluValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrIle 397
Db 1696 TTGCTCGAGAGAGTGAAGAAACAGCTACATTTGTGTAACCTTAGAGCAACATTCATC 1755
Qy 398 AsnGlyAspLeuPheAlaValPheAla 406
Db 1756 AATGCTGATATCT 1782
RESULT 12
AX069359 1971 bp DNA linear PAT 25-JAN-2001
LOCUS AX069359
DEFINITION Sequence 23 from Patent WO0102600.
ACCESSION AX069359
VERSION AX069359.1 GI:12579222
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Yuan,C.S.
TITLE Detection of analytes using attenuated enzymes
JOURNAL Patent: WO 0102600-A 23 11-JAN-2001;
GENERAL ATOMICS (US)
FEATURES
source location/Qualifiers
1..1971


```

Db      847 GAGGTGATGAGCAATCAAGACAAGATGCTGCCATCGCCAACTATGCGATGAGCTG 906
Qy      81 G|YthrluMetCysLysLysLleLeuAlaHisGly---LleLysThrLeuHisLleuTyr 99
Db      907 GCCGTGAGCTGTGGCAGAGAGCTTGTGGCACTGGCTGGTGCAGAGCTTCCTCACTTCTAC 966
Qy      100 ThrLeuAsnMetGluLysSerSerAlaLeuAlaLleLeuMetAsnLeuGlyLleuLleuGlu 119
Db      967 ACCCTCAACCGCGAGATGGCTACCAAGAGTGTGAGAGCCCTGGGGATGTGACACTGAG 1026
Qy      120 SerLysValSerArgSerLeuProTrrArgArgProAlaAsnValPheArgValLysGlu 139
Db      1027 GACCC---AGGGTCCCTTACCTGGGCTCTCACTGGCCACCCCAAGCCCGAGAGGAA 1083
Qy      140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrLleSerArgThrIleGly 159
Db      1084 GATGTAGCTCCCATCTTCTGGGCTCTCAGACCAAGATTAACNTACCGTACCAGAG 1143
Qy      160 TrpAspLlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db      1144 TGGAGACAGATTCCCTAACGCGCGCTGGGCAATTCTTCCCTGGCTTGGGAGACTG 1203
Qy      180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
Db      1204 AAGGACTACTACCTCTTCACTGAGAGCAAGTCCCCCAAGAGAGAGCTGTAGATG 1263
Qy      198 TrpAlaValProLeuLysSerValGluAspLleTyrGluArgPheArgLeuTyrCysLeu 217
Db      1264 TGGGGGAGAGAGCTGACAGTCAAGCAAGTCTTGTGAAGCTTGTCTTACCTCTCG 1323
Qy      218 G|YLysLeuArgSerAsn-----ProTrrSerGluLeuAspGly 230
Db      1324 GGAAGAACCAACCGGAATGGTCAACAAGAGCTTCCCTGCCCTGGAAGAT---GAGCC 1380
Qy      231 LeuGlnProGluTyrLysLleLleAsnGluGluLeuGlyLysLleAsnThrLysGlyPhe 250
Db      1381 CTGGCGGCTGAGACAGCTCTGAGAGAGCTGTGGCGGGTGAACCCCGCAGGCGCATC 1440
Qy      251 LeuThrLleAsnSerGlnProAlaValAsnGlyLysSerSerAspSerProThrValGly 270
Db      1441 CTCACACATCACTCAACGCCCAACATCAACGGGAAGCGCTCTCCAGCCCATGCTGGC 1500
Qy      271 TrpGlyGlyProGlyLysTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db      1501 TGGGGGCCCCAGCGGGGCTATGTCTTCCAGAGAGGCTTATAGATTTTCACTTCCCGC 1560
Qy      291 G|ULysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
Db      1561 GAGACAGCGGAACACTTCTGCAGAGTGCAGAAAGTACAGAGCTCGGGTTAATTACAC 1620
Qy      310 AlaValAsnLysAspGlySerTrrLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db      1621 CTGTGCATATGTGAAGGATGAATAACATCAACAATGCCCTCCAAATGCACAGCCGAAATGCTGTC 1680
Qy      330 ThrTrpGlyValPheProAlaLysGluLleLeuGlnProThrLleValAspProValSer 349
Db      1681 ACTTGTGGGATCTTCCCTGGGAGAGATCATCCAGCCACCTGATAGTCCCGTACG 1740
Qy      350 PheAsnValTrrLysAspLysAlaPheGluLleTrrPheSerArgGlyTrrAlaLysLeuTyr 369
Db      1741 TTATGTCTTGGAGAGACAGGCTTTTCCCTGTGGATTTGAGCGGTGGGAAAGCTGTAT 1800
Qy      370 ProGluAspGluLysAspLysLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
Db      1801 GAGGAGAGAGTCCCGTCCCGCAGCATTCACAGTACATCCAGCAGCACTACTTCTGTGTC 1860
Qy      390 SerLeuValAspAsnAspTyr---LleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db      1861 AACCTGTGTGACATGACTTCCACTGGACAACTGCTGTGGAGTGTGGAGAGAC 1917

```

```

RESULT 14
AX050441
LOCUS      AX050441      2187 bp      DNA      linear      PAT 12-JAN-2001

```

```

DEFINITION      Sequence 1 from Patent WO0071754.
ACCESSION      AX050441
VERSION      AX050441.1      GI:12226648
KEYWORDS
SOURCE
ORGANISM      human.
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Johnson, W.G., and Stenroos, E.S.
TITLE      Methods for diagnosing, preventing, and treating developmental
                disorders due to a combination of genetic and environmental factors
                University of Medicine and Dentistry of New Jersey (US)
                location/Qualifiers
                1=2187
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT      479 a      655 c      616 g      437 t
ORIGIN
Alignment Scores:
Pred. No.:      5 5e-77      Length:      2187
Score:      979.00      Matches:      188
Percent Similarity:      63.728      Conservative:      79
Best Local Similarity:      44.878      Mismatches:      138
Query Match:      44.764      Indels:      14
                Gaps:      7
US-09-720-451-6 (1-408) x AX050441 (1-2187)
Qy      1 AlaArgValLleValThrGlnLeuPheTyrAspTrpThrAspLlePheLeuLysPheValAsn 20
Db      667 GCCGATTTCATCAATCAAGCAGAGCTTCTTGTGAAGCTGACACATCTTCCGCTTGTGTAAG 726
Qy      21 AspCysArgGlnLleGlyLleThrCysProLleValProGlyLleMetProLleAsnAsn 40
Db      727 GCATGCACACCAATGGGCACTACTGCCCATGCTCCCGGATCTTCCATGCACAGGC 786
Qy      41 TyrLysGlyPheLleArgMetThrGlyPheCysLysThrLysLleProAlaAspLleMet 60
Db      787 TACCACTCCCTTGGCGAGCTGTGAAAGCTGTCCAAAGCTGGAGGTGGSACAGAGAGAG 846
Qy      61 AlaLleGluLysProLleLysAspAsnGluLysAlaValLysAlaTyrGlyLleLysLeu 80
Db      847 GAGCTGATGAGCAATCAAGACAAGATGCTGCATCGCCATGAGCACTATGCGATGAGACTG 906
Qy      81 G|YthrluMetCysLysLysLleLeuAlaHisGly---LleLysThrLeuHisLleuTyr 99
Db      907 GCCGTGAGCTGTGGCAGAGAGCTTGTGGCACTGGCTGGTGCAGAGCTTCCTCACTTCTAC 966
Qy      100 ThrLeuAsnMetGluLysSerSerAlaLeuAlaLleLeuMetAsnLeuGlyLleuLleuGlu 119
Db      967 ACCCTCAACCGCGAGATGGCTACCAAGAGTGTGAGAGCCCTGGGGATGTGACACTGAG 1026
Qy      120 SerLysValSerArgSerLeuProTrrArgArgProAlaAsnValPheArgValLysGlu 139
Db      1027 GACCC---AGGGTCCCTTACCTGGGCTCTCACTGGCCACCCCAAGCCCGAGAGGAA 1083
Qy      140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrLleSerArgThrIleGly 159
Db      1084 GATGTAGCTCCCATCTTCTGGGCTCTCAGACCAAGATTAACNTACCGTACCAGAG 1143
Qy      160 TrpAspLlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db      1144 TGGAGACAGATTCCCTAACGCGCGCTGGGCAATTCTTCCCTGGCTTGGGAGACTG 1203
Qy      180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
Db      1204 AAGGACTACTACCTTCTCACTGAGAGCAAGTCCCCCAAGAGAGAGCTGTAGATG 1263
Qy      198 TrpAlaValProLeuLysSerValGluAspLleTyrGluArgPheArgLeuTyrCysLeu 217

```

[illegible]

| | | | |
|---|---|---------------|------|
| JOURNAL | Submitted (17-MAY-1994) Rima Rozen, Pediatrics, Human Genetics and Biology, McGill University - Montreal Children's Hospital, 2300 Tupper St., Montreal, Quebec H3H 1P3, Canada | | |
| REFERENCE | 4 (bases 1 to 2196) | | |
| AUTHORS | Leclerc,D., Sibani,S. and Rozen,R. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (29-OCT-1999) Pediatrics, Human Genetics and Biology, McGill University - Montreal Children's Hospital, 2300 Tupper St., Montreal, Quebec H3H 1P3, Canada | | |
| REMARK | Amino acid sequence update by submitter | | |
| COMMENT | On Nov 2, 1999 this sequence version replaced gi:945022. | | |
| FEATURES | | | |
| source | 1. 2196 /organism="synthetic construct" /db_xref="taxon:32630" | | |
| gene | 1. 2196 /gene="MTHFR" | | |
| misc_feature | 1. . 8 /gene="MTHFR" /note="EcoRI linker" 13. .1983 /gene="MTHFR" /note="Method: conceptual translation with partial peptide sequencing" /codon_start=1 /transl_table=1 /product="methyltetrahydrofolate reductase" /protein_id="AA44440.2" /db_xref="GI:6139053" /translation="MVNENRGNSSINPCLRESASSGSSKSDSRGCTPGIDPRHREHREMRRLRSGDKWFLEPPPTATGAVNLIIRPRMAAGELYIDVYHVRADGDSIDKTSMMIASTAVNYCGELTILHMTCCRLEITGHLKAKQLGKNTMALRGDPIDGWEEEGEFNAYVDLVKIRSEFDYDTCVAGYPKGPEGSEADIKLKEKVSAGADITLTOLTFEADPTFFRFVAKCTMGITCIPVGIPIQGHSLRQIVKLSKLEVPOELKVDIPEIKDNDAAIRNGLIAVLSICELASGLVGHFTYLNREMAETEVLKRLGWGTDPRRLDIPALSAHKRREEDVRPIFMASSRKSYYTQOENDEPNRGWSSLSRPAWGTGLKDYLYFLYLSKSPKELLMMCEELTLESAVEFVLYLISGPNRNGHKVYTCIPMNDPIPLAETSLIKBELLRVNRGILITINQPINKRPSSDPIYGMPSGGYVFOKALEFTSRKETAALLQVLYKLYELRNWTLVWKGNTINAEIOPNATWIFGFEKRIIQTPTVDPVSEFMFKDEAFALMIIRNMGWLYEESPSRTIIOYIHNNFLVNLVDNAP"DEPDDNCIMQVVEDTLLELNRPONARETAP" | | |
| variation | 1298 /gene="MTHFR" /note="E429A; reported by Weisberg et al (1998) Mol. Genet. Metabol. 64, 169-172" /replace="a" 2196 /gene="MTHFR" /note="18 A nucleotides" /evidence="experimental" | | |
| polya_site | | | |
| BASE COUNT | 482 a 657 c 618 g 439 t | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 5.53e-77 | Length: | 2196 |
| Score: | 979.00 | Matches: | 188 |
| Percent Similarity: | 63.72% | Conservative: | 79 |
| Best Local Similarity: | 44.87% | Mismatches: | 138 |
| Query Match: | 44.76% | Indels: | 14 |
| DB: | 12 | Gaps: | 7 |
| US-09-720-451-6 (1-408) x HSU09806 (1-2196) | | | |
| QY | 1 AIAATGVALIIIEVALThrcInleuPheryrAspIInrAspIlePheluIuysPheValaSn 20 | | |
| DB | 676 GCCGATTTCATCATCAGCAGGCTTTTCTTGAGCGTACACACATCTTCCGCTTGTGAAG 735 | | |
| QY | 21 ASpcysArgGInIIeGlyIleThrcysProIleValProGlyIYleMetProIleasnSn 40 | | |
| DB | 736 GCATGCACCGACATGGGACTTCATGCCCCCAACGTGTCCCGGGATCTTTCCATCCAGGCG 795 | | |
| QY | 41 TTYrLySgLyPheIleArgMetThrgLyPhecLySvTfThLySIIleProAlaaspIleMet 60 | | |

Db 796 TACCACTCCCTTGGCAGCTGTGAAGCTGTCCAAAGCTGGAGGTGCCACAGAGATCAAG 855
QY 61 AlaAlaLeuGluProIleuValAspAsnGluGluAlaValAlaValAlaValIleuHisLeu 80
Db 856 GACCTGATGTGAGCAATCAAGACACGATGCTGCATCCGACATCTTGGCATCGAGCTG 915
QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr 99
Db 916 GCCGTGAGCCTGTGTCAGAGAGCTTCTGCGCAGTGGCTTGGTCCAGGCTCCACTTCTAC 975
QY 100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu 119
Db 976 ACCCTCAACCGCGAGAGGTACACACAGAGTGTCTGAAGCGCTGGGATGTGACTGAG 1035
QY 120 SerLysValSerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu 139
Db 1036 GACCCC---AGCGTCCCTTACCTGGGCTCTCAGTGCACCCCAAGCGCGAGAGAA 1092
QY 140 AspValArgProIlePheTyrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
Db 1093 GATGTACGTCCCATCTTCTGGGCTCCAGACCAAGATTACATCACTACCTACCCAGAG 1152
QY 160 TyrAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
Db 1153 TGGGACGAGTTCCTTAACGCGCTGGGCAATCTCTCCCTGCTGGGGAGCTG 1212
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
Db 1213 AAGGACTACTACTCTCTTACTGTAAGAGCAAGTCCCCCAAGAGAGAGCTGTAAGATG 1272
QY 198 TyrAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
Db 1273 TGGGGGAGGAGGCTGACACAGCAAGTGTCTTGAAGTCTTGTCTTACTCTCG 1332
QY 218 GlyLysLeuArgSerAsn-----ProThrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCGGAATGCTCACAAAGTACTTGCCTGCTGGAAAGAT--GAGCCC 1389
QY 231 LeuGlnProGluThrLysIleIleAsnGluGluLeuGluLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGCGGCTGAGACAGCCTGTGTAAGAGAGAGCTGTGGGGTGAACCGCCAGGCAATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly 270
Db 1450 CTCACCATCACTACACAGCCACATCAACGGGAGCCGCTCCGACCCCATCGTGGGC 1509
QY 271 TyrGlyGlyProGlyTyrValTyrGluLysAlaTyrValGluPheGlySerLys 290
Db 1510 TGGGGCCCCGCGGGGCTATGCTCTTCCAGAAAGCCTACTTAGATTTTTCACTTCCGC 1569
QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp--ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGGAGACTTCTGCAAGTGTGAAGATACAGAGCTCCGGTTAATTACAC 1629
QY 310 AlaValAsnLysAspGlySerTyrPlysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTGTCAATGTGAAGGTGAAACATCACCAAGCCCTGAACCTGCACCGCAATGCTGTC 1689
QY 330 ThrTyrGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGGCATCTTCCCTGGGCGAGAGATCATCCAGCCCACTAGTGCATCCCGTCAGC 1749
QY 350 PheAsnValTyrPlysAspGluAlaPheGluIleTyrSerArgGlyTyrPalaSerLeuTyr 369
Db 1750 TTCATGTCTTGGAAAGACGAGCGCTTGCCTGTGATTGAGCGGTGGGAAAGCTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluValGlyGlySerHisPheLeuVal 389
Db 1810 GAGGAGGAGTCCCGCTCCGACCATCATCCAGTACATCCACGACAACTACTCTGTC 1869
QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
Db 1870 AACCTGTGAGACATGACTTCCCACTGACAACTGCTTGGCAGGTGTTGGAAGAC 1926

Search completed: February 15, 2003, 00:16:56
Job time : 3670 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 12, 2003, 17:03:09 ; Search time 58 Seconds

(without alignments)
2157.312 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVITOLFVDTDFLEKFN.....VSLVDNDYINGDLFAVADF 408

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_plus_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09720451/runat_05022003_074809_2384/app_query.fasta_1.583
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=rni -MIMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR-ext -HEAPSIZE=500 -MNLBN=0 -MAXLEN=200000000
-USBR=US09720451 @cgn1.1.17 @runat_05022003_074809_2384 -NCPD=6 -ICPD=3
-NO_XLPHY -NO_MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database:

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 979 | 44.8 | 1971 | 4 | US-09-347-878-23 |
| 2 | 979 | 44.8 | 2187 | 4 | US-09-318-448-1 |
| 3 | 979 | 44.8 | 2219 | 3 | US-08-738-000-3 |
| 4 | 979 | 44.8 | 2219 | 4 | US-09-258-928-3 |
| 5 | 979 | 44.8 | 2220 | 3 | US-08-738-000-1 |
| 6 | 979 | 44.8 | 2220 | 4 | US-09-258-928-1 |
| 7 | 148.5 | 6.8 | 1449 | 4 | US-08-858-207A-54 |
| 8 | 148.5 | 6.8 | 28882 | 4 | US-08-961-527-140 |
| 9 | 138 | 6.3 | 4848 | 4 | US-08-951-957A-1 |
| 10 | 138 | 6.3 | 4848 | 4 | US-08-951-957A-4 |
| 11 | 138 | 6.3 | 4848 | 4 | US-08-951-957A-6 |
| 12 | 109.5 | 5.0 | 1556 | 2 | US-08-687-355A-3 |

| | | | | | | |
|----|-------|-----|-------|---|--------------------|--------------------|
| 13 | 109.5 | 5.0 | 1556 | 4 | US-09-407-367-3 | Sequence 3, Appli |
| 14 | 104.5 | 4.8 | 14311 | 4 | US-08-646-695-1 | Sequence 1, Appli |
| 15 | 104.5 | 4.8 | 14311 | 4 | US-08-646-695-7 | Sequence 7, Appli |
| 16 | 104.5 | 4.8 | 14311 | 5 | PCT-US96-06053-1 | Sequence 1, Appli |
| 17 | 104.5 | 4.8 | 14311 | 5 | PCT-US96-06053-7 | Sequence 7, Appli |
| 18 | 100 | 4.6 | 38584 | 4 | US-09-453-702B-50 | Sequence 50, Appli |
| 19 | 97.5 | 4.5 | 8220 | 2 | US-08-568-459A-11 | Sequence 11, Appli |
| 20 | 97.5 | 4.5 | 8220 | 2 | US-08-487-826B-11 | Sequence 11, Appli |
| 21 | 97.5 | 4.5 | 8220 | 4 | US-09-210-288-11 | Sequence 11, Appli |
| 22 | 97.5 | 4.5 | 19124 | 2 | US-08-487-826B-13 | Sequence 13, Appli |
| 23 | 93.5 | 4.3 | 1900 | 4 | US-08-965-762-22 | Sequence 22, Appli |
| 24 | 93.5 | 4.3 | 1900 | 4 | US-09-911-927-22 | Sequence 22, Appli |
| 25 | 93.5 | 4.3 | 1900 | 4 | US-09-911-927-24 | Sequence 24, Appli |
| 26 | 93.5 | 4.3 | 1900 | 4 | US-09-911-882-22 | Sequence 22, Appli |
| 27 | 93.5 | 4.3 | 1900 | 4 | US-09-911-882-24 | Sequence 24, Appli |
| 28 | 93 | 4.3 | 2800 | 4 | US-09-056-285A-7 | Sequence 7, Appli |
| 29 | 92.5 | 4.2 | 1654 | 4 | US-09-221-017B-326 | Sequence 326, App |
| 30 | 89 | 4.1 | 1366 | 1 | US-08-670-354-5 | Sequence 5, Appli |
| 31 | 89 | 4.1 | 1366 | 4 | US-09-320-424-5 | Sequence 5, Appli |
| 32 | 89 | 4.1 | 1366 | 5 | PCT-US96-10895-5 | Sequence 5, Appli |
| 33 | 89 | 4.1 | 2144 | 4 | US-08-876-798A-1 | Sequence 1, Appli |
| 34 | 89 | 4.1 | 2144 | 4 | US-08-876-798A-3 | Sequence 3, Appli |
| 35 | 87.5 | 4.0 | 1320 | 1 | US-08-290-448A-75 | Sequence 75, Appli |
| 36 | 87.5 | 4.0 | 1320 | 1 | US-08-290-448A-75 | Sequence 75, Appli |
| 37 | 87.5 | 4.0 | 1320 | 1 | US-08-175-069A-75 | Sequence 75, Appli |
| 38 | 87.5 | 4.0 | 1320 | 4 | US-08-461-939B-75 | Sequence 75, Appli |
| 39 | 87.5 | 4.0 | 1320 | 4 | US-08-461-939B-75 | Sequence 75, Appli |
| 40 | 87.5 | 4.0 | 1320 | 4 | US-09-393-529-1 | Sequence 1, Appli |
| 41 | 87.5 | 4.0 | 2835 | 1 | US-08-750-532-2 | Sequence 2, Appli |
| 42 | 87.5 | 4.0 | 4765 | 1 | US-08-750-532-8 | Sequence 8, Appli |
| 43 | 87.5 | 4.0 | 4765 | 4 | US-08-894-818B-7 | Sequence 7, Appli |
| 44 | 87.5 | 4.0 | 4765 | 4 | US-09-445-472-5 | Sequence 5, Appli |
| 45 | 87 | 4.0 | 3619 | 4 | US-08-377-503-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-09-347-878-23
: Sequence 23, Application US/09347878C
: Patent No. 6376210
: GENERAL INFORMATION:
: APPLICANT: Iuon, Chong
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
: FILE REFERENCE: 25885-1651
: CURRENT APPLICATION NUMBER: US/09/347,878C
: CURRENT FILING DATE: 1999-07-06
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 23
: LENGTH: 1971
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Human methylenetetrahydrofolate reductase (MTHFR)
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: AF105977/genbank 1-11
US-09-347-878-23

Alignment Scores:

Pred. No.: 7,42e-110
Score: 979.00
Best Local Similarity: 63.728
Best Local Similarity: 44.87%
Query Match: 44.87%
DB: 4
Gaps: 7

US-09-720-451-6 (1-408) x US-09-347-878-23 (1-1971)

QY 1 AAlaryVallleValhrgInleuPhetYrAsphrAspleleuYsPheValaAn 20
Db 664 GCGATTCATCAACACGACGCTTTCTTGTAGCGTCAACATCTCCGCTTGAGAG 723

| | | | |
|----|------|--|------|
| QY | 21 | ASPCYARGINLLEGLYLIETHCHSPROILEVALPROGLYIILEMETHPROILEANAN | 40 |
| DB | 724 | GCATGCACCGACATGGCATCTGCGCCATCGTCCCGGGATCTTCCCATCCAGGCC | 783 |
| QY | 41 | TYRLSGLYPHEILEARYMETIRGLYPHECYLSYTHIRLYSILEPROILAAPILMET | 60 |
| DB | 784 | TACCACTCCCTTCGGCAGCATCTGTCAACCTGTCCAGCTGGAGGTGCCACGAGGATCAAG | 843 |
| QY | 61 | ALALALEUGLIPROILEYASPAISNGULIALVALIALYSALATYRGYLIETHLEU | 80 |
| DB | 844 | GAGGTGATGGACCAATCAAGAACAGAGGTGGCCATCCGCACTTGGCATCGAGCTG | 903 |
| QY | 81 | GLYTHIRGUMETCYSLYSLYLIIELEUALHISGLY---IIELYSTHIRHISLEUTYR | 99 |
| DB | 904 | GCCGTAGCCCTGTGCACGAGACCTTGTGGCCAGTGTGGTGCCAGGCGCTCCACCTTCAC | 963 |
| QY | 100 | THIRLEUASNMETGLYSERSEALALEUALILEMETASLEUAGLYLEULIEGLUGL | 119 |
| DB | 964 | ACCCCTCAACCGGAGAAATGGCTACACACGAGGTGCTGAAGCCCTGGGGATGTGACTGAG | 1022 |
| QY | 120 | SETLYVALSERASERSETLEPROTPRAGAAGYPROLOLAANVALPHEATYVALYSGL | 139 |
| DB | 1024 | GACCCC---AGCGGTCCCTACCTGGCGCTCTCAGTGGCCACCCAGCCGCGAGAGGA | 1081 |
| QY | 140 | ASPVALAARGPROILEPHEPTRIPALASNAAGPROLYSERYTIIIESEARATHIRILEG | 159 |
| DB | 1081 | GAGTACGTCGTCATCTTGTGGCGCTCCACGACCAAGAGTTCATCTCCGATCCCGAGG | 1140 |
| QY | 160 | TRPASPGLINTYRPROHISGLYATYRTPGLYASPSERYCAANPROSETRYRGLYALALEU | 179 |
| DB | 1141 | TGGGAGAGATTCCTTCMAAGCGCGCTGGGCGCATCTCTCCCTCGCTTGGGAGAGCG | 1200 |
| QY | 180 | SERASTYRGLN-----PHEMAYRPROAGYALAAAGSPYLSYLSLEUVALIGLUGL | 197 |
| DB | 1201 | AAGGACTTACTACCTTCTTACCTGAAGAGCAANTCCCCAGAGAGAGCTGCTGAAGAG | 1267 |
| QY | 198 | TRPALVALPROLEULYSEVALIGUASPILETYRGLUAGPHEATYRGLYCYSEU | 217 |
| DB | 1261 | TGGGGGAGGAGCTACCCAGCAAGCAAGTCTTGGAAAGTCTTGTTCCTTACCTCTCG | 1320 |
| QY | 218 | GLYLYLEUAYSERASN-----PROTPSERGLUEUAASGLY | 230 |
| DB | 1321 | GGAGAACCAACCGGAATGGTCACAAAGTGCACCTGCGCCCTGAGACGAT--GAGCCC | 1377 |
| QY | 231 | LEUGLIPROGLINTYRILEILEANISGLUGLISLEUVALYSLEANTHRYSLYRPE | 250 |
| DB | 1378 | CTGGCGGCTGAGACAGCGCTCTGAGAGAGAGAGTCTGCGGTGAACCCGCAAGGCATC | 1437 |
| QY | 251 | LEUTHIRLEANSERGLNPROIALVALAENGLYGLYLSERASPSERPROTHVALGILY | 270 |
| DB | 1438 | CTCAACATCAATCTCACGCCCAACATCAACGGGAGCCGTCTCCGACCCCACTGGGGC | 1497 |
| QY | 271 | TRPGLYLPYRGLYGLYTYRVALTYRGLINLYSALATYRVALIGLUPHEPACYSERYLS | 290 |
| DB | 1498 | TGGGGCCCAAGGGGGGCTATGTTCTCCAGAAAGGCTACTAGAGTTTTCACCTCCCG | 1557 |
| QY | 291 | GLULYLEUASALALEUVALASPLYSGLYSASP---ARYTHSERILEUTHIRYRMET | 309 |
| DB | 1558 | GAAKACGCGGAAGCACTTCTCCAGAGTGTGAAGAACTACAGCTCCGGGTAAATTCAC | 1617 |
| QY | 310 | ALAVALASNLVASPSGLYSETRTPYLSERASVALIGLYGINTHIRASPVALAANALVAL | 329 |
| DB | 1618 | CTTGTCAAATGTAAAGGAGGAAACATCACCAATGCCCCCTGAACGACGCGATGCTGTC | 1677 |
| QY | 330 | THIRTRIGLYVALPHEPROIALVALSGULIILEGLINPROTHIRILEVALASPPROVALSER | 349 |
| DB | 1678 | ACTTGGGCGCATCTTCCCTGGGCGAGAGATCATCCAGCCACCGATAGTGATCCCGTAC | 1733 |
| QY | 350 | PHEASNLVALTPYLYASPSGLUALAPHEGLIETRPSERATRGLYTRPALASERLEUTYR | 369 |
| DB | 1738 | TTATATCTTCTGAGAGAGAGGCGCTTGTGCTGTGATGTAGGCGGTGGGAAAGCTGTAT | 1797 |

```

QY 370 ProGluAspGluAlaISerArgIysLeuValGluGluValGlyGlySerHisPheLeuVal 389
      |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 1798 GAGGAGAGGAGCCCGCGGACCGACCATATCATCCAGTATCATCGACGACAACTATCTTGGGCG 1857
QY 390 SerLeuValAspAsnAspTyr--LeasnGlyAspLeuPheAlaValAlaPheAsp 407
      ::|||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 1858 AACCTGATGGACAAAGACTTCCACTGAGACACTGCTCTGGCAGGTGTGGAGAAC 1914

RESULT 2
US-09-318-448-1
: Sequence 1, Application US/09318448
: Patent No. 6210950
: GENERAL INFORMATION:
: APPLICANT: Johnson, William G.
: TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
: TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
: FILE REFERENCE: 601-1-057
: CURRENT APPLICATION NUMBER: US/09/318,448
: NUMBER OF SEQ ID NOS: 46
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 2187
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-318-448-1

Alignment Scores:
Pred. No.: 8.8e-110 Length: 2187
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 4 Gaps: 7

US-09-720-451-6 (1-408) x US-09-318-448-1 (1-2187)
QY 1 AlaArgValIleValIleThgInLeuPheTyrAspThrAspIlePheLeuIysPheValAsn 20
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 667 GCGGATTTGATCATCATGACCCACCTTTCTTTGAGGCTGACACATTTCTCCGCTTTGTGAAG 726
QY 21 AspGcArgGlnIleGlyIleThcYsProIleValProGlyIleMetProIleAsn 40
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 727 GCATGACGCGACATGCGCATCTTGCCGCGCATGCGCCCGGAGATCTTCCATCCAGGCGC 786
QY 41 TyrIleGlyPheIleArgMetThcGlyPheCysIysThrIysIleProIleAspIleMet 60
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 787 TACACTCCCTTGGCGACCTTGAGAGCTGTCCAGCTGAGAGTGCCACGAGGACTAAG 846
QY 61 AlaAlaLeuGluProIleIysAspAsnGluGluAlaValIysAlaTyrGlyIleHisLeu 80
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 847 GAGCATATTGACCCGATCAAGACAGCATGTCGCCATCCGCAACTTTCGATCGACGCTG 906
QY 81 GlyThrGlnMetCysIysIleLeuAlaHisGly--IleIysThrHisLeuIleTyr 99
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 907 GCGGTAGCGTGTGCCAGGACTCTTGCGCAGTGGCTGTGTCGACAGGCTCCACTTTCAC 966
QY 100 ThrLeuAsnMetGlyIysSerAlaLeuAlaIleLeuMetAsnLeuGlyIleGlu 119
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 967 ACCCTCAACCGGAGATGGCTACACAGAGTGTGAAGCGCCGTGGGAGTGTGACTGAG 1026
QY 120 SerIysValSerArgSerLeuProTyrPargArgProIleAsnValIlePheArgValIysGlu 139
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 1027 GAGCCC--AGCGTCGCCCTACCTCGTGGGCTGTCAATGCGCACGCCAACGCCGAGAGGAA 10833
QY 140 AspValArgProIlePheThrAlaAsnArgProIysSerTyrIleSerArgThrIleGly 159
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 1084 GAGTGTAGCGTCCATCTTGTGGCGCTCAAGCAACAAAGATTCATCTACCTTACCCAGAGG 11433
QY 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyIleLeu 179
      ||| |||... |||... |||... |||... |||... |||... |||... |||... |||...
Db 1144 TGGGAGAGATTTCCCTTAAGCGCCGCTGGGAGATTTCTTCCCTCGGCTTTGGGAGGCTG 12033

```


| | | | |
|----|------|---|------|
| QY | 330 | THTPGPIYALPHIEPCOHALYSGIILLETLEGINPOTHRIETVALASPROVALSER | 343 |
| | | | |
| | 1690 | ACCTGGGGCACTTCTCCTGGGGCAGAGATCATCCAGCCACCGTAGTGATCCCGTCAGC | 1749 |
| Db | | | |
| QY | 350 | PHASNVALTRPFLYASPGIUALAPHGIIUETSPERATGGLYTRPPALASERLEUYR | 369 |
| | | | |
| | 1750 | TTCATGCTTCTGGAAAGACAGACGAGCGCTTGTCCCTGTGTGATGTAGCGCGTGGGAAACGTGAT | 1809 |
| Db | | | |
| QY | 370 | PROGIIASPGIUALASERATGRLYSLEUALGIUUALGLYLSERTHISPHLEUAL | 389 |
| | | | |
| | 1810 | GAGGAGAGATCCCGTCCCGCACCATCATCCAGTGCATCCACGACCAACTACTTCCGTGTC | 1869 |
| Db | | | |
| QY | 390 | SEIIEUALASPAASAPYR---LIEASNGIYASPLEUPHEIALVALPHEIALASP | 407 |
| | | | |
| | 1870 | AACCTGTGGGCAATGACTTCCCATCGGACCACTCGCTTCCGCGAGGTGGGAAGAC | 1926 |
| Db | | | |

RESULT 5
US-08-738-000-1

; GENERAL INFORMATION:

TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE REDUCTASE

NUMBER OF SENTENCES: 74

NUMBER OF SEQUENCES. 1
CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: Continental Plaza - 411 Hackensack Avenue

; CITY: Hackensack

STATE: New Jersey

COUNTRY: U.S.A.
ZIP: 07601

COMPUTER READABLE
ZLF: 07001

MEDIUM TYPE: FLOPPY

COMPUTER: IBM PC compati

OPERATING SYSTEM: PC-DOS/MS

SOFTWARE: PALENTIN Release
CURRENT APPLICATION DATA:

CONTRACT DESCRIPTION DATA:
APPLICATION NUMBER: US/08/738,000

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION ;

APPLICATION NUMBER: WO PCT/CA95/00314
FILING DATE: 35-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9410620.0

FILING DATE: 26-MAY-

```

; INFORMATION FOR SEQ ID NO: 1:

```

```

; SEQUENCE CHARACTERISTICS:
;
;     LENGTH: 3330 base pairs

```

LENGTH: 2220 base pairs
TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genom)

HYPOTHETICAL: NO

```

; AN11-SENSE: NO
;
FEATURE:

```

| NAME/KEY: | CDS |
|-----------|-----|
| PHONE: | |

LOCATION: 1..1980

US-08-738-000-1

1

Alignment Scores: 0.030-1

Score: 979 00
Plea: NO.: 9:02e-1

Percent Similarity: 63.72%

Best Local Similarity: 44.878

| | | | |
|--------------|--------|---------|----|
| Query Match: | 44.76% | Indels: | 14 |
| DB: | 3 | Gaps: | 7 |

US-09-720-451-6 (1-408) x US-08-738-000-1 (1-2220)

QY 1 ALaRgVAlIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
 ||| |||::||| |||::: ||| ||| ::||| |||
 Db 676 GCCGATTTCATCATCAGCAGAGCTTTCTTTTGAGGCTGACACATCTTCCGCTTTGGCAG 735

| | | | |
|----|------|---|------|
| OY | 21 | ASPCATGATglnllieglllylthrhcyprollevalprogllymetprolleamsn | 40 |
| Db | 736 | GCATGACCCAGACATGGGATCTACTGGCCCATCGCCCGGATCTTCCATCAGAGCC | 795 |
| OY | 41 | TYTfysgLyPheIleArgmetThrIglyPheCysLySfThrIlysIleProAlaAspIleMet | 60 |
| Db | 796 | TACCACTCCCTTCGGACAGCTTGGAAGCTGTGCACAGCTGAGGAGTGCAGAGATCAAG | 855 |
| OY | 61 | AlAlAlaLeuGlnProIleIlyAspAsnGluGlnAlaValIySAlaTyrglyIleHisLeu | 80 |
| Db | 856 | GACGTATTGACCCATCAAGAACAGATGCGCATCCCAACTGTGGCATCGAGCTG | 915 |
| OY | 81 | gLyThrlmEcCysLySfIlySfIleAlaIleHisgLy--IleYsThrlmHisLeuTy | 99 |
| Db | 916 | GGCGTAGCCTGTGCACAGACCTTGGCAGAGCGCTTGGACAGGCTCCACTTCAC | 975 |
| OY | 100 | ThrlmAsnMetGlnIlySserAlaIleAlaIleIleMetAsnLeuGlyLeuIleGluI | 119 |
| Db | 976 | ACCCCTAACCCGAGATGCGATCCACACAGAGTGTGAAGCCCTGGGGATGTGACTAG | 1035 |
| OY | 120 | SerLyValSerArgSerLeuProTyArgArgProAlaAsnValPheArgValIysGlu | 139 |
| Db | 1036 | GACCC--AGCGGCCCTCCCTGGGCTGTCAGGCCACCCCAAGCCGAGAGNA | 1092 |
| OY | 140 | AspValArgProIlePheThrlAlaAsnArgProIlySserTyrlIleSerArgThrlIegly | 159 |
| Db | 1093 | GATGTACGTCCCATCTTGTGGCTCCAGACCAAGACTACATCTACGCTATCCAGAG | 1155 |
| OY | 160 | TrpAspGlnTyProHisgLyIArgTyrglyAspSerCysAsnProSerTyrglyAlaLeu | 179 |
| Db | 1153 | TGGGAGAGTATCCCTTAAGGCGCGTGGGCAATCTCTCCCTGGGAGCTG | 1212 |
| OY | 180 | SerAspTyrgln-----PheMetArgProArgAlaIArgAspLySfIlyValIleGluI | 197 |
| Db | 1213 | AAAGACTACTACCTTCTTACTGTAGAGCAAGTCCCCCAAGAGAGCTGCTGAAGATG | 1272 |
| OY | 198 | TrpAlaValProLeuIlySserValGluAspIleTyrglyArgPheArgLeuTyrcysLeu | 217 |
| Db | 1273 | TGGGGAGAGACTGACCATCCAGAGCAAGTGTCTTGAAGCTTGTCTTACCTCTCG | 1332 |
| OY | 218 | GlyLyLeuAlArgSerAsn-----ProTpsSerGluAlaAspGly | 230 |
| Db | 1333 | GGAGAACCAACCGGATGCTGCACAAAGTACTTGCCTGCCCTGGAACAT--GAGCCC | 1388 |
| OY | 231 | LeuGlnProGlnThrLysIleIleAsnGluGlnLeuGlnIlySfIleAsnThrLysGlyPhe | 250 |
| Db | 1390 | CTGGCGCGTGAAGACAGCTGCTGAGAGAGAGCTGCGCGGTGAACCCCGAGGCATC | 1445 |
| OY | 251 | LeuThrIleAsnSerGlnProAlaValAsnGlyIlySserAspSerProThrValIgly | 270 |
| Db | 1450 | CTCAACATCAATCAACACCCCAACATCAAGAGGAAAGCCGTCTCGAACCCATCGTGGCC | 1505 |
| OY | 271 | TrpGlyLyProGlyLyTyValTyrglnIlySAlaTyrglyValGluPhePheCysSerIys | 290 |
| Db | 1510 | TGGGAGCCAGCGGGGCTATGCTTCCACAAAGCGCTACTTGAAGTTTTCACCTCCCG | 1565 |
| OY | 291 | GlnIlySLeuAspAlaIleValAspLyCysIlyAsp--ArgThrSerLeuThrTyrmec | 309 |
| Db | 1570 | GAGACAGCGGAACACTTCTCAAGTGTCTGAAGAAAGTACAGCTCCGGTTAATTCAC | 1628 |
| OY | 310 | AlAlaValAsnIlyAspGlySerTyrlPlySserAsnValGlyGlnThrAspValAsnAlaVal | 329 |
| Db | 1630 | CTTGCTCAATGTGAAGGTGAAGAAACATCAACCAATGCCCTCGAACCTGCAGCGAATGCTGTC | 1689 |
| OY | 330 | ThrlTrpGlyValPheProAlaIlySglnIleIleGlnProThrIleValAspProValSer | 349 |
| Db | 1690 | ACTTGGGAGCATCTTCCCTGGGCGAGAAATATCAACGCCACGTAAGTATCCGCTGAC | 1749 |
| OY | 350 | PheAsnValTyrlPlySAspGluAlaPheGlnIleTyrlPserArgGlyTyrlPAlaSerLeuTy | 369 |
| Db | 1750 | TTCTATGCTCTGGAAGAGAGAGCCCTTTCGCCCTGTGATGAGCGGTGGGAAGCTGTAT | 1809 |

QY 370 PROUAAEPGLNALASerarglysleuvalglugluvalglglglssrthspheulval 389
 1810 GAGAGAGATCCCGCGCCGACCATCATCTCATCTCATCATCAGCAACATCTTCCTGGTC 1869
 QY 390 Serleuvalasphanaspyr--11easnglyaspheupheulvalphealaasp 407
 1870 AACCTGTGTGAGAAATGACTTCCACATGAGCAACTCCCTCTGGCAGGTGGTGAAGAC 1926

```

RESULT 6
US-09-258-928-1
Sequence 1, Application US/09258928
Patent No. 6218120
GENERAL INFORMATION:
APPLICANT: ROZEN, Rima
APPLICANT: GOYTEE, Philippe
TITLE OF INVENTION: GNA FOR HUMAN METHYLENETETRAHYDROFOLATE
TITLE OF INVENTION: REDUCTASE
FILE REFERENCE: 04644/005002
CURRENT APPLICATION NUMBER: US/09/258,928
CURRENT FILING DATE: 1999-03-01
PRIORITY APPLICATION NUMBER: 08/738,000
PRIORITY FILING DATE: 1997-02-12
PRIORITY APPLICATION NUMBER: GB 9410620.0
PRIORITY FILING DATE: 1994-05-26
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2220
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1980)
US-09-258-928-1

```

| | |
|------------------------|------------|
| Alignment Scores: | |
| Pred. No.: | 9, 02e-110 |
| Score: | 979.00 |
| Percent Similarity: | 63.72% |
| Best Local Similarity: | 44.87% |
| Query Match: | 44.76% |
| DB: | 4 |
| Gaps: | 7 |
| Length: | 2222 |
| Matches: | 188 |
| Conservative: | 79 |
| Mismatches: | 138 |
| Indels: | 14 |
| Gaps: | 7 |

US-09-720-451-6 (1-408) x US-09-258-928-1 (1-2220)

| | | | |
|----|------|---|------|
| QY | 1 | AlaAaGValIleValThrGlnLeuPheMetYaaPheAspIlePheLeuLysPheValAsn | 20 |
| | | | |
| Db | 676 | GCGCAATTCATCATCAGCAGCACTTCTCTTGGAGCTGACACATCTCTCCGTTTTGGAG | 735 |
| QY | 21 | AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn | 40 |
| | | | |
| Db | 736 | GATGCGACGAGATGGGATCACTGTGCCCATCGTCCCGGAGCTTGTCCATCCAGGAC | 795 |
| QY | 41 | TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMet | 60 |
| | | | |
| Db | 796 | TACACATCCCTTCGGCAACTGTGGAGCTGTCCAGCTGTCCAGCTGCACAGAGATCAAG | 855 |
| QY | 61 | AlaAlaLeuGluProIleLysAspAsnGluAlaValLysAlaTyrGlyIleHisLeu | 80 |
| | | | |
| Db | 856 | GAGCTGATATGACCAATCAAAAGCAACAGATGCTGCATCCCAACTGTGCATCAGAGCG | 915 |
| QY | 81 | GlyThrGluMetCysLysIleLeuAlaLysIle---IleLysThrHisLysLeuTyr | 99 |
| | | | |
| Db | 916 | GCGGTAGCTGTGCACAGAGCTTGTGGCAGAGCGCTTGTCACAGCGCTCCACTGTAC | 975 |
| QY | 100 | ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu | 119 |
| | | | |
| Db | 976 | ACCTTACACCGCGAATGGCTACCAACAGAGTGTCTAAAGCCCTCGGGGAATGTGACTGAG | 1035 |
| QY | 120 | SerLysValSerArgSerLeuProTyrPargArgProAlaAsnValPheArgValLysGlu | 139 |
| | | | |
| Db | 1036 | GAGCCCC---AGCGTGGTCCCTTACCTGTGGGTCTACATGGTCAACCCAGACGCGAGAGA | 1092 |

| | | | |
|----|------|--|------|
| QY | 140 | AspValAlaArgProIlePheThrPrAlaAsnArgProIlySserTryIleSerArgThrIleGly | 159 |
| Db | 1093 | GAAGTACGCGCCCATTTCTTGGGCTCCAGACCAAGAGATTTCATCTTCACCTACCAGAG | 1152 |
| QY | 160 | TrpAspGlnTyrProHleGlyAlaGlyTrpGlyAspSerCysAsnProSeryTrgAlaLeu | 179 |
| Db | 1153 | TGGGAGAGATTCCCAAGAGGGCGGTGGGGATCTCTCTCTCCCTGGCTTTGGGGAGCTG | 1212 |
| QY | 180 | SerAspTyrGln-----PheMetArgProAlaArgAlaArgAspIlySleValGluGlu | 197 |
| Db | 1213 | AAAGACTACTACTCTTCTTACTCTCAAGACAGAGTCCCCCAAGGAGAGCTCGTGAAGATG | 1272 |
| QY | 198 | TrpAlaValProIleuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu | 217 |
| Db | 1273 | TGGGGGAGAGACTCAGCAGAGGAACAGAGTCTTGAAGTCTTTGTTTACCTTCG | 1332 |
| QY | 218 | GlyLysLeuArgSerAsn-----ProTrpSerGluLeuAspGly | 230 |
| Db | 1333 | GGAGAACCAACCGGAGATGTCACAAAGTACACTGCTCCCTCTCGAACAT--GAGCCC | 1389 |
| QY | 231 | LeuGlnProGluThrLysIleLeuAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe | 250 |
| Db | 1390 | CTGGCGGCTGAGACAGAGCTCTGTAAGAGAGAGCTGTCGGGTGAACCGCAGGCAATC | 1449 |
| QY | 251 | LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly | 270 |
| Db | 1450 | CTACACATCAATCTCACACCCCAATCAACAGGAGAGCCGTCTCCGACCCCATGTGGGCT | 1509 |
| QY | 271 | TrpGlyGlyProGlyGlyTyrValTyrValTyrGlnLysAlaTyrValGluPhePheCysLeu | 290 |
| Db | 1510 | TGGGGCCCAAGCGGGGCTATGCTTCCAGAGAGCCCTACTTGAAGTWTTCACCTTCCGC | 1569 |
| QY | 291 | GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet | 309 |
| Db | 1570 | GAGACACGGAGAACTTCTTCAAGTCTCTCAAGAGAGTACAGAGCTCCGGTTAATTAACAC | 1629 |
| QY | 310 | AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal | 329 |
| Db | 1630 | CTTGTCATATGTAAAGGAGAAATCATCACCATGCGCCCTGAAGTCAGCGAGTCTGTGC | 1689 |
| QY | 330 | ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer | 349 |
| Db | 1690 | ACTTGGGGCATCTTCCCTGGGCGAGAGATATCTACACCCACCGTAGTGGATCCCGTACG | 1749 |
| QY | 350 | PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTrpAlaSerLeuTyr | 369 |
| Db | 1750 | TTCATGTTCTGCAGAGAGAGGCTTTCGCCCTGTGATGTGACGCTGGGGAAGCTGTAT | 1809 |
| QY | 370 | ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerThrPheLeuVal | 389 |
| Db | 1810 | GAGAGAGAGTCCCTCCCGACCATATCTACAGTATCTACAGCAACATCTTCCCTGGTGC | 1869 |
| QY | 390 | SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp | 407 |
| Db | 1870 | AACTCTGTGACAAATGACTTCCCACTGGACACCTGCTCTTCTGGACGTTGGAGAGAC | 1926 |

RESULT 7
US-08-858-207A-54
Sequence 54, Application US/08858207A
Patent No. 6348328
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
TITLE OF INVENTION: No. 6348328e1 Compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA

1 COUNTRY: USA
 2 ZIP: 19406-0939
 3
 4 COMPUTER READABLE FORM:
 5 MEDIUM TYPE: Diskette
 6 COMPUTER: IBM Compatible
 7 OPERATING SYSTEM: DOS
 8 SOFTWARE: FastSeq for Windows Vers
 9
 10 CURRENT APPLICATION DATA:
 11 APPLICATION NUMBER: US/08/858,207A
 12 FILING DATE: 09-MAY-1997
 13 CLASSIFICATION: 435
 14 PRIOR APPLICATION DATA:
 15 APPLICATION NUMBER: 60/017670
 16 FILING DATE: 14-MAY-1996
 17 ATTORNEY/AGENT INFORMATION:
 18 NAME: Gimmil, Edward R.
 19 REGISTRATION NUMBER: 38,891
 20 REFERENCE/DOCKET NUMBER: P50475
 21 TELECOMMUNICATION INFORMATION:
 22 TELEPHONE: 610-270-4478
 23 TELEFAX: 610-270-5090
 24
 25 TELEX:
 26
 27 INFORMATION FOR SEQ ID NO: 54:
 28
 29 SEQUENCE CHARACTERISTICS:
 30 LENGTH: 1449 base pairs
 31 TYPE: nucleic acid
 32 STRANDEDNESS: single
 33 TOPOLOGY: linear
 34
 35 US-08-858-207A-54

| | | |
|------------------------|----------|------------|
| Alignment Scores: | | |
| Pred. No.: | 3,326-08 | 1449 |
| Score: | 148.50 | 45 |
| Percent Similarity: | 49.65% | 26 |
| Best Local Similarity: | 31.47% | 57 |
| Query Match: | 6.79% | 16 |
| DB: | 4 | Gaps: 5 |

```

0Y      4  ILeValhIhrGlnLeuPheTyraSphThrAsp11IlePheLeuLysPheValAsnAspCysArg 23
Db      870  CTCCTTAACCTAGCCTCTTCTTGCACATGAGCCGCTTCATGATTTTTCAGAGCAAGCTATT 929
0Y      24  GlnIleGlyIlePheTrcCysProIleValProGlyIleMetProIleAsnAspTyrLysGly 43
Db      930  TTGGCTGGGATTGATGTCCCATTTCACTTCATGCGAGGATTTATGCCATTCGTGAATCGAATCAG 989
0Y      44  PheIleArgMetThrGlyPheCysLysThr---LysIleProAlaAspIleMetAlaAla 62
Db      990  GCCTCCGACACTCTTGAAGACTCTGTGGAATATCCATCTTCCACGCCAAATTTAAAGCCATC 1049
0Y      63  LeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThr 82
Db      1050  TTAAGCAAGATGATGACATGACATGCCCTGAGTGCCTGACGACGACGAGACTTGCCATGCAATG 1109
0Y      83  GluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsn 102
Db      1110  GACCAAAATCGTAGCATTGGTCACTCAAGATGTTGCCGCTGCATCTCTATATCGATGAAT 1169
0Y      103  MetGluLysSerAlaLeuAlaIle-----LeuMetAsn----- 113
Db      1170  AATGCTGATACAGCAAAATACATCCATCAAGCAACCCATGCTTGTTTATATCCAGCT 1229
0Y      114  LeuGlyLeuIleGlu-----GluSerLysValSerArgSerLeuPro-----Tyr 128
Db      1230  CTAGAGATG-ATATAAAGCAAAACCATCTTCTTCAGGTGAGGGGAGATGGTCTTTTAAATGG 1288
0Y      129  ArgArgPro 131
Db      1289  CAAAGACT 1297

```

```

US-08-961-527-140/c
: Sequence 140, Application US/08961527
: Patent No. 6420135
:
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
:
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
:
: INFORMATION FOR SEQ ID NO: 140:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28882 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
US-08-961-527-140

```

| Accession | Score | Percent Similarity | Best Local Similarity | Query Match | DB |
|---|----------|--------------------|-----------------------|-------------|----|
| US-09-720-451-6 (1-408) x US-08-961-527-140 (1-26882) | 4,52e-06 | 148.50 | 52.38% | 6.79% | 4 |
| QY 4 IlevaIthrGlnLeuphetyrAspThrAspIlePheLeuLysPheValAsnAspCysAsn 23 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| DB 15802 CTCGCACTCAGCGCTCTTTCGACATGAGCCCTCTTGATTTTCGAGCAAGTGATAC 15743 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| QY 24 GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsnTyLysGly 43 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| DB 15742 TTGGCTGGGATTGATGTGCCATTATCGAGGAATTAAGCCAAATTCGATCGAAATCG 15683 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| QY 44 PheIleGlyMetThrGlyPheCysThr--LysIleProAlaAspIleMetAlaAla 62 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| DB 15682 GCTCCGACCTCTTGAGAACTGTGAGATATCCATCTTCCACGCCAAATTAAAGCCATC 15623 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| QY 63 LeuGluProIleLysAspAsnGluGluAlaValLysAlaTyLysIleHisLeuGlyThr 82 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| DB 15622 TTAGCAAGATGATGCATGACCTGCTGCTCAAGCAAGCAAGACTGCCATGCAGTC 15566 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| QY 83 GluMetCysLysLysIleLeuAlaHisGlyIleTyThrLeuHisLeuTyThrLeuAsn 102 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| DB 15566 GACCAAAATCGTGAGCTTGATTAAGTCAAGGATGTGCCGGGATGCAATCTGATACGATGAAT 15503 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |
| QY 103 MetGluLysSerAla 107 | 148.50 | 52.38% | 31.43% | 6.79% | 4 |

Db 15502 AATGCTGATACAGCA 15488

RESULT 9

US-08-955-957A-1

Sequence 1, Application US/08955957A

Patent No. 6312920

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Rostock Jr., Paul R.

TITLE OF INVENTION: SAM Operon

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,957A

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: P-10162

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4848 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 986..2209

NAME/KEY: CDS

LOCATION: 2241..3341

US-08-955-957A-1

Alignment Scores:

Pred. No.: 4.68e-06 Length: 4848

Score: 138.00 Matches: 30

Percent Similarity: 43.64% Conservative: 18

Best Local Similarity: 27.27% Mismatches: 62

Query Match: 6.31% Indels: 0

DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-1 (1-4848)

OY 5 ValThrgInleuPhetYrAspThrAspIlePheLeuLysPheValAsnAspCysArgIn 24

DB 3872 ATCAGCGAGATGTTCTTCGCGGTCCGAGACTATCTGCGGTCGCGGACCGGGTCGCGCC 3931

OY 25 IleGlyIleThrcysProIleValProGlyIleMetProIleAsnAspTyrLysGlyPhe 44

DB 3932 GCCGCGCTGTCGACCCGCGTATCCCGGATCATGCCCGCACCGAGCTCGCGCATC 3931

OY 45 IleArgMetThrnglyPheCysLysThrLysIleProIleAspIleMetAlaIalaLeuGlu 64

DB 3992 GCGCGCTTCGCGAGCTTTCACGCGCACCTTCCCGAGAGGTCTCGCGCGCGCGTGGAG 4051

OY 65 ProIleLysAspAsnGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMet 84

Db 4052 GCCCGCCGCGCACCGCGCGAGGAGACCGCATCGGGGTGAGTACCGCCCGCATG 4111

OY 85 CysLysIleIleAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsnMetGlu 104

DB 4112 GCCGCGCGCTCTCTCGCGAGAGGCCCGCGGACATGCTACATCATCTCAACCGGCTCC 4171

OY 105 LysSerAlaLeuAlaIleLeuMetAsnLeu 114

DB 4172 ACCGCGACGTGTGAGATCCACCGGAGACATC 4201

RESULT 10

US-08-955-957A-4

Sequence 4, Application US/08955957A

Patent No. 6312920

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Rostock Jr., Paul R.

TITLE OF INVENTION: SAM Operon

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,957A

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: P-10162

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4848 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 3338..4255

US-08-955-957A-4

Alignment Scores:

Pred. No.: 4.68e-06 Length: 4848

Score: 138.00 Matches: 30

Percent Similarity: 43.64% Conservative: 18

Best Local Similarity: 27.27% Mismatches: 62

Query Match: 6.31% Indels: 0

DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-4 (1-4848)

OY 5 ValThrgInleuPhetYrAspThrAspIlePheLeuLysPheValAsnAspCysArgIn 24

DB 3872 ATCAGCGAGATGTTCTTCGCGGTCCGAGACTATCTGCGGTCGCGGACCGGGTCGCGCC 3931

OY 25 IleGlyIleThrcysProIleValProGlyIleMetProIleAsnAspTyrLysGlyPhe 44

DB 3932 GCCGCGCTGTCGACCCGCGTATCCCGGATCATGCCCGCACCGAGCTCGCGCATC 3931

```

Oy      45  ILeIrgvethrghlyPhecySLySThrlYsllelProAlAsPlleMeAlAlaleuSlu 64
          |||
Db      3992 GCGGCGTCCGCCAGACTTTTCCACGGCACCTTCCCAGAAgGTCTCGCGCGGCGCTGGAg 4051
          |||
Oy      65  ProlleLysaspasnngluAlalaVallylsAlatYrglylleHslleugSlYThrlglumet 84
          :::: |||
Db      4052 GCGCGCGCGGCAACCGCGCGGAGGAGACAACGCATTCGGGGGTGGAATTACGCCACCOCCTATG 4111
          :::: |||
Oy      85  CyslYslYslleleAlahIslglylleYslThrlleuHslleuYTrhleuSnsMetgu 104
          :::: |||
Db      4112 GCCCGCGCGGTCTCGCGCGGAGGCGCGCGCGGAGCTGCACTACATCATCCTCTCAACCGCTCC 4171
          :::: |||
Oy      105  LysSerAlaleuAlalAleuMetasleu 114
Db      4172 ACCGCGACTTGAGATTCACCGGAACATC 4201

RESULT 11
US-08-955-957A-6
; Sequence 6, Application US/08955957A
; Patent No. 6312920
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Rosteck Jr., Paul R.
TITLE OF INVENTION: SAM Operon
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,957A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: P-10162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHEtical: NO
ANTI-SENSE: NO
US-08-955-957A-6

Alignment Scores:
Pred. No.: 4.68e-06 Length: 4848
Score: 138.00 Matches: 30
Percent Similarity: 43.64% Conservative: 18
Best Local Similarity: 27.27% Mismatches: 62
Query Match: 6.31% Indels: 0
DB: 4 Gaps: 0

US-09-720-451-6 (1-408) x US-08-955-957A-6 (1-4848)
Oy      5  ValThrGlnleuPheryrrAsPlnAsPllePheleuLysPhevalAsmspcysArgrGln 24
          :::: |||
Db      3872 AUCACGAGAGUUCUUCGCGGUCGAGAGACUAUCUGCGGUCGCGGAGCCGGUGUCGCGGCC 3931

```

```
QY      25   IleglylEthrcysProillevAlProglyllEmePrOIlAsmaAnstnyLrGsylyphe 44  
          |||           |-----|||-----|||       :--:  
Db     3932 GCCGCCTCCACCGCCCCGGATCACCAGGAATAAUCGCCGCCACCGCATCCUGCGGCAGAUGC 3991  
  
QY      45   lIeargMethTrhgLyPheCyLSyThTyLSleProalaspIlemetaLaaleuglu 64  
          |||         |||             |||        ::    ||||||  
Db     3992 GCGGGCUUCCGCGAGCUUUUCCCAGCACCUUCCCGCAAGGUUDCGCGCGGTUGGAG 4051  
  
QY      65   ProillysaspaNgJngJuallavalYSalaTYrglylleHisLeuglyThrGlumet 84  
          :::::|||               |||:::      |||  |||  
Db     4052 GCCGCCCGCGGCACAACCGCGCGAGGACACCGCANUCGGGUGAGUAGACGCCCGCCAUG 4111  
  
QY      85   CyslSySLileualaHslGyllyleLTyrThrhIsleutyTrrhreusnmetglu 104  
          ::::::::::::||||||  |||||  |||||  |||||  |||||  
Db     4112 GCCCGCGCGCTCCCGCGCGGAAAGGGGCCCGCGGATCGACAUCAUCAUCUCCAACGCCUCC 4171  
  
QY      105  LysSerlaIleaualalleumeLasneu 114  
          ::::|||  |||  |||:::  
Db     4172 ACCGCGAGCUUGAGUAUCCACCGGAACAUC 4201  
  
RESULT 12  
US-08-687-355A-3  
; Sequence 3, Application US/08687355A  
; Patent No. 5989834  
; GENERAL INFORMATION:  
APPLICANT: Synaptic Pharmaceutical Corporation  
TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
TITLE OF INVENTION: Y/PEPTIDE YY (YY) RECEPTORS AND USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
  
COMPUTER READABLE FORM:  
MEDIA TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MIS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,355A  
FILING DATE: NO. 5989834ember 26, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPW/MAT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NOS: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1556 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOCHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 211..1353  
US-08-687-355A-3
```

| | |
|---------------------------------|------------------------------|
| Alignment Scores: | |
| Pred. No.: 0.00227 | Length: 1556 |
| Score: 109.50 | Matches: 81 |
| Percent Similarity: 35.42% | Conservative: 49 |
| Best Local Similarity: 22.07% | Mismatches: 136 |
| Query Match: 5.01% | Indels: 101 |
| Dt: 2 | Gaps: 20 |


```

QY 128 Trp-----ArgArgProAlaAsnValPheArgValLysGlu----- 139
Db 727 TGGGGTGCAGCGCTGCTGCTGGCAAGTCCCTTCCCATCTTCGGGAGTACTACTATT 786
QY 140 AspValArgProIlePhe-----TrrPalaAsnArgProLys 151
Db 787 GAGATTATTCTCGACTTTCAGATTGTAGCCCTGTACTGAGAAATGGCCGGGAGAGAGAG 846
QY 152 Ser-----TyrIleSerArgThr-----Ile 158
Db 847 AGTGTGTACGGGTACAGTCTACAGCCTTTCACCTGCTAATCCTTACGTTTTCCTGTG 906
QY 159 GlyTrrPargInTyrProHisGlyArg---TrrPglYasSerCysAsn---ProSerTyr 176
Db 907 GGCATCATATCTTCTCTCCACCCGGATCTGGAGTAAGCTAAAGAACACACGTTAGTCT 966
QY 177 GlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGlu 196
Db 967 GAGAGTGCAGAGTACCATTTACCATCAGCGAAGGACAAAGACCAAAATGCTCGTG--- 1023
QY 197 GluTrrPalaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCys 216
Db 1024 -----TGC 1026
QY 217 LeuGlyLysLeuArgSerAsnProTrrPserGluLeuAspGlyLeuGlnProGluThrLys 236
Db 1027 GTGGTAGTGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
QY 237 IleIleAsnGluGlnLeuGlnLysIleAsnThrLysGlyPheLeuThrIleAsnSerGln 256
Db 1087 ATCGACAGCAGCATGCTCGTACCTGAGAGGATCAAAACATCATTCACCGTGTCCACATT 1146
QY 257 ProAlaValAsnGlyLysSerAspSerProThrValGlyTrrPglYglProGlyGly 276
Db 1147 ATTCGATGTGCTCCACCTTCCCAACCCCTTCTGTATGCTGATGATGACAGCAAC--- 1203
QY 277 TyrValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeu 296
Db 1204 -----TACGAAAGAGCTTCTCTGACGCTTCGCTGAGCAGAGGTGATGCTCAT 1257
QY 297 ValAspLysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspLysSer 316
Db 1258 -----CACTCGAGAGGTGCTCATGACCTTCAAGGCT---AAAAAGAACCTGGA 1302
QY 317 TrrPlySerAsnValGlyGlnThrAsp-----ValAsnAla----- 328
Db 1303 GTCAAAAGAACACATGGCTCCTGCTCTTTTTCAGAGGCCACCAACGCTGTAAGATGC 1362
QY 329 -----ValThrTrp-----GlyValPhePro 335
Db 1363 TGTGAAGATACGTGGTAATTGCGACAGAGTGGCCAACTGGTTAGGGAAGGTT--- 1419
QY 336 AlaLysGluIleIleGlnProThrIleValAspProValSerPheAsnValTrrPlyAsp 355
Db 1420 TCTGGCTAGTGCATGCCACCTCCATTGATGTGACCCCTAAAGACCAAGATGGGAAG 1479
QY 356 GluAlaPheGluIleTrrPserArgLysTrrPalaSerLeuTyrProGluAspGluAsp 375
Db 1480 CACGGATATTGTTCTCTGGAAACTGGCTGGAGAAAT-----GAGGAGAAATAAAC 1530
QY 376 ArgLysLeuValGluGluVal 382
Db 1531 AGATTGCTGTGGCGCAACGTT 1551

```

```

RESULT 14
US-08-646-695-1
; Sequence 1, Application US/08646695
; Patent No. 6189943
; GENERAL INFORMATION:
; APPLICANT: Rose, John K.
; TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
; NUMBER OF SEQUENCES: 44

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,695
FILING DATE: On Even Date Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 760..2025
FEATURE:
NAME/KEY: CDS
LOCATION: 2092..2886
FEATURE:
NAME/KEY: CDS
LOCATION: 2946..3632
FEATURE:
NAME/KEY: CDS
LOCATION: 3774..5306
FEATURE:
NAME/KEY: CDS
LOCATION: 5429..11755
US-08-646-695-1
Alignment Scores:
Pred. No.: 0.356 Length: 14311
Score: 104.50 Matches: 71
Percent Similarity: 34.06% Conservative: 54
Best Local Similarity: 19.35% Mismatches: 117
Query Match: 4.78% Indels: 125
Gaps: 17
US-09-720-451-6 (1-408) x US-08-646-695-1 (1-14311)
QY 80 LeuGlyTrrPglInGlyMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyr 99
Db 691 ATGGGACGAGACAAACAAACATATTATCATTAATAAGCCTGACGAGAACTTAAACA 750
QY 100 ThrLeuAsnMetGlyLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGlu 119
Db 751 GTAATCAAAATGCTGTGTTACAGTCAAGATCATTTGACACACAGTCAATGATGTTCCAAA 810
QY 120 SerLysValSerArgSerLeuProTrrPargArgProAlaAsnValPheArgValLysGlu 139
Db 811 CTTCCTGCAAAATGAGAGAT---CCAGTGAATATACCGGACGATTTACTTACGAAATGCAAG 867
QY 140 AspValArgProIlePheTrrPalaAsnArgProLysSer-----Tyr 153
Db 868 GAGATT---CCTCTTTAC---ATCAATACTACAAAAGCTTTGTGATGATCTAAGAGATAT 921

```



```
Db 13207 G T A C T T C C A G A T G A G T T A C G G A T C T T C C A G A A C C A G C G A G A T G A C A A A T G C T T G C C T 13148
      ::::
QY 211 -----ArgPheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTrpSerGlu--- 227
      ||||| ::::| ::::| ::::| ::::| ::::|
Db 13147 T T G T A T C T A C T T G G C T T A T C A G A T G G G C A G A C A A A A T G C C T G A A T A C A G A A A A A A G 13088
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 228 ---LeuAspGlyLeuGlnProGluThrLysIleIleAsnGluGlnLeuGlyLysIleAsn 246
      ::::| ::::| ::::| ::::| ::::| ::::|
Db 13087 C T C A T G A T G G G C T C A C A A A T C A A T G A C A A A A A T G A T G C A A C A C T T G A A C C T C T T G T G 13028
      ::::| ::::| ::::| ::::| ::::| ::::|
QY 247 ThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSer 266
      ::::| ::::| ::::| ::::| ::::|
Db 13027 C C A G A A G G T ----- 13010
      ::::| ::::| ::::| ::::| ::::|
QY 267 ProThrValGlyTrpGlyLysProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhe 286
      ||| ||||| ||| ||| ::::|
Db 13009 T T T G A T G T G ---T G G G A A A T G A C A G T A A T T A C A C A A A A A T T G T C G C T G C A G T G G A C A T G 12953
      ::::| ::::| ::::| ::::| ::::|
QY 287 Phe-----CysSerLysGluLysLeuAspAlaLeuValAsp 298
      ||| ::::| ::::| ::::| ::::|
Db 12952 T T C T T C C A C A T G T C A A A A A C A T G A A T G T G C C T G T C A G A T A C G A C T A T T G T T C C 12893
      ::::| ::::| ::::| ::::| ::::|
QY 299 LysCysLysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTrpLys 318
      ::::| ::::| ::::| ::::|
Db 12892 A G A T T C A A A G A T T G C T G C A T T G G C A C A C A T T -----G G A C A C C T C T G C 12848
      ::::| ::::| ::::| ::::|
QY 319 SerAsnValGlyGlnThrAspValAsnAlaValThrTrp----- 331
      ||| ::::| ::::| ::::|
Db 12847 A A A A T A A C C G A A T G C T A C A G A A G A T G T A A C G A C C T G A C T T G A A C C A G A A G T T G C A 12788
      ::::| ::::| ::::| ::::|
QY 332 -----GlyAlaPheProAlaLysGluIle-----IleGln 341
      ::::| ::::| ::::| ::::|
Db 12787 G A T G A A A T G G T C C A A A T G A T G C T T C C A G G C C A A G A A A T T G A C A G A G C G A T C A T A C A T G 12728
      ::::| ::::| ::::| ::::|
QY 342 ProThrIleValAsp-----ProValSerPheAsnValTrpLysAspGlu 356
      ||| ::::| ::::| ::::|
Db 12727 C C T T A T T G A T G A C A C T T T G A T T G T C T T A A G T C C A T A T T C T T C G C T A A A A A C C C T 12668
      ::::| ::::| ::::| ::::|
QY 357 AlaPheGluIleTrpSerArg 363
      ||||| ||| ::::|
Db 12667 G C C T T C C A C T T C T G G G G C C A A 12647
      ::::| ::::| ::::|
```

Search completed: February 12, 2003, 17:06:31
Job time : 112 secs


```

      PRIOR APPLICATION NUMBER: GB 9410620.0
      PRIOR FILING DATE: 1994-05-26
      NUMBER OF SEQ ID NOS: 14
      SOFTWARE: FASTSEQ for Windows Version 4.0
      SEQ ID NO 2
      LENGTH: 660
      TYPE: PRT
      ORGANISM: Homo sapiens
      US-09-258-928-2

Query Match      44.8%; Score 979; DB 4; Length 660;
Best Local Similarity 44.9%; Pred. No. 5,1e-91;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps

OY      1 ARVVTQLEFDTGFIKFFVNDRCROIGITCPYVGPMPINNKKGFIRMTGCKTKIPADIM 60
DB      226 ADFTITDLEFADTFRRFVACCTDNGTCITCPYVGFPIQGVHSLROLVKLSLEVPQELK 285

OY      61 AALEPIKDNEAVKAYGILHGTBCKKIIAHG-ITLHLITLNNKESALMIINMGLIEE 119
DB      286 DVIPRIDNDAAIRNNGIELAVSLCOELLASGLVGLHPTLNNEMATEVLRKLGWTE 345

OY      120 SKVSRSLPWRRPANNVFRKEDVPIPMANRPKSYISRTIGNDQYPRHGWGDSNPSYAL 179
DB      346 DP-RRLPLWALSAHPRRREDVPIPMASRPKSYIIRYQWEDDEFNGWSSSPARGEL 404

OY      180 SDYO--EMRPARDKKLVEEAVPLKSVEDIYERRLICLGKLSN-----PWSLGD 230
DB      405 KDYVLEFLTKSPKEELIKMMWGSELTSEASVFEVVLXLSGEPNNGHKVTCLPMND-EP 463

OY      231 LQPEFKIINDELEKINKGFLITNSOPAVNGKSDSPYVGMGPGGYVYOKAYVEFCSK 290
DB      464 LAEESLKEELRLVNNQGITLITNSQPNINGKRPSSDPYVGMGPGSGYVFOKAYLEFFTSR 523

OY      291 EKLDALYDKCD-RTSLTYMAVNRKDSMKSNNVQGDVNAVTVGCFPAKEIIQPIITVDS 349
DB      524 EFAELALQVLAKKYLRLVNYHLVWKEBNTNAPQLPNAVTVGIPGREIITQPIVVDVS 583

OY      350 ENWKDEDFELWMSGMSLXPRDEDAARKLYEEGSHFVSLVYNDY--INGDLFAVFAD 407
DB      584 FFWKDEAFALMIERWGLKLYEESPSRTIIQIHDNYFLVNLVNDPFLDCLMQLQVYED 642

RESULT 6
US-08-858-207A-324
; Sequence 324, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670

```


Db 274 TYFHATNCRGERTKGYCRNDQVPTV---FDYVPOYL-----WFEEMADFCRRKN 323
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNPW----- 225
Db 324 KKIKDKVRCRCRKEDKEDKDRYCSRNGYDCEKTKRAIGKLCISCLYACNPFYDWIN 383
QY 226 SELDGLQPETKILNEQLEKINTKGFLLTNSQPAVNGEKSD---SPTVGMGPGGYVYQKA 282
Db 384 NKEQGFQKKKKYDEIKKYE-----NGASGSRQKRDAGGTTTNYDG-----YEKK 431
QY 283 Y-----VEFGCKEKLDALVDCKDRT---SLTYAVVNDGSMKSNVGQTDVNAV 329
Db 432 FYDELNKSEYRTVDFLEKLSNEEICTKVKDEEGTIDFKNVSD---STSGASGTNVE 487
QY 330 TWGVFPAKEIIOP-----TIYDVSFNVWKDEAFETWSRGWASLY---PEDEASRKL 379
Db 488 SGTTFYRSKYCPCPCYCGVKKVNNGSSNMEERKNCKCKSG--KLYEPKPREGTTITI 545
QY 380 EEVGGSHFLVSLVDN---DYINGD 400
Db 546 LKSGKHDIIEKLNKFCDEKNGD 569

RESULT 9

US-08-487-826B-12
Sequence 12, Application US/08487826B

GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORGANISM: Plasmodium falciparum
US-08-487-826B-12

Query Match

4.5%, Score 97.5; DB 2; Length 2710;

Best Local Similarity 19.8%, Pred. No. 3.6;
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADYMAAL-EPKNDENAVKAYGHLGTEM---CKLIANGIKTLHYLTNMRKSLALIL 111
Db 182 PSQICMTLARSFADIGDIYRGDIYLGNDPEIKORQOLENNLKTI----- 226
QY 112 MNLIEESKVSRSLLPWRPNAVFRKEDVRFIFW-ANRPKSYISRTIGMDQYPRGRMGD 170
Db 227 --FGKITYELKNGAEARYGNDPEFFKLRED---WMTANR-----ELV-WKATLCNMG 273
QY 171 S-----CN-----PSYCALSDY--QTRPRARKKUIVEWAVPL----- 202
Db 274 TYFHATNCRGERTKGYCRNDQVPTV---FDYVPOYL-----WFEEMADFCRRKN 323
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNPW----- 225
Db 324 KKIKDKVRCRCRKEDKEDKDRYCSRNGYDCEKTKRAIGKLCISCLYACNPFYDWIN 383
QY 226 SELDGLQPETKILNEQLEKINTKGFLLTNSQPAVNGEKSD---SPTVGMGPGGYVYQKA 282
Db 384 NKEQGFQKKKKYDEIKKYE-----NGASGSRQKRDAGGTTTNYDG-----YEKK 431
QY 283 Y-----VEFGCKEKLDALVDCKDRT---SLTYAVVNDGSMKSNVGQTDVNAV 329
Db 432 FYDELNKSEYRTVDFLEKLSNEEICTKVKDEEGTIDFKNVSD---STSGASGTNVE 487
QY 330 TWGVFPAKEIIOP-----TIYDVSFNVWKDEAFETWSRGWASLY---PEDEASRKL 379
Db 488 SGTTFYRSKYCPCPCYCGVKKVNNGSSNMEERKNCKCKSG--KLYEPKPREGTTITI 545
QY 380 EEVGGSHFLVSLVDN---DYINGD 400
Db 546 LKSGKHDIIEKLNKFCDEKNGD 569

RESULT 10

US-09-210-288-12
Sequence 12, Application US/09210288

GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2710 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHEICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Plasmodium falciparum
;; US-09-210-288-12

Query Match 4.5%; Score 97.5; DB 4; Length 2710;
Best Local Similarity 19.8%; Pred. No. 3.6;
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMALL-EPIDNEAVKAVGIHIGTEM---CKKILAHGKTLHLTYLNMKESALAIL 111
DB 182 PSQICTMLANSFADIDYVGRDLYGNPQETIKQKQLENNLKITI----- 226
QY 112 MNLGLIEESKYSRSLPWRPRANVRYKEDVRPIFW-ANRPKSYISRTIGMDQYHGRMGD 170
DB 227 --PEKIYEKLNAGARVGNDEFFKLRD-----WWTANR-----ETV-WKAITCNAMGN 273
QY 171 S-----CN-----PSYGALSDY--QEMPRAROKKILEEVAVPL----- 202
DB 274 TYFNATCRGERTKGYCRMDQVPTY---FDIVQYLR-----WFEWAEDFCCKKN 323
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNFW----- 225
DB 324 KRIDVNRNCRGKDEKDKRYCSRNGYDCEKTKRAIGKRYGKOCISCLYACNRYVDMIN 383
QY 226 SELGLOPETRIINEOLEKINTGFLTINSOPAVNGEKSD---SPYVGSGPGGYVYOKA 282
DB 384 NQKQFPQKQKKYDEIKYTE-----NGASGSRQKRDAGTITTYDQ-----YEKK 431
QY 283 Y-----VEFCSKEKLDALYDKCKDRT--SLTYMAVNDGSKSVNGQTDVNAV 329
DB 432 FYDELINSEKRYVDKFLKLSNEICTKYKDEGGTIDFKVNSD---STSGASGTNVE 487
QY 330 TWGVFPKKEIIOF-----TIYDPSVFNWKRDEAFELMSRGMASLY---PEDEASRKL 379
DB 488 SQGTFYRSKCYQCPGYGVKRVNNGSSNMEWEKKNNGCKSG--KLYPRKDKGTTITTI 545
QY 380 EGVGSHFLVSLVDN---DYINGD 400
DB 546 LKSGGHDIIEKLNKFCDEKNGD 569

RESULT 11
US-08-487-826B-14
; Sequence 14, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chluis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/487,826B
;; FILING DATE: 10-SEP-1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Israel, Ned
;; REGISTRATION NUMBER: 29,655
;; REFERENCE/DOCKET NUMBER: NH121.001CPI
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 235-8550
;; TELEFAX: (619) 235-0176
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3060 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-487-826B-14

Query Match 4.5%; Score 97.5; DB 2; Length 3060;
Best Local Similarity 19.8%; Pred. No. 4.3;
Matches 88; Conservative 59; Mismatches 142; Indels 155; Gaps 27;

QY 56 PADIMALL-EPIDNEAVKAVGIHIGTEM---CKKILAHGKTLHLTYLNMKESALAIL 111
DB 180 PSQICTMLANSFADIDYVGRDLYGNPQETIKQKQLENNLKITI----- 224
QY 112 MNLGLIEESKYSRSLPWRPRANVRYKEDVRPIFW-ANRPKSYISRTIGMDQYHGRMGD 170
DB 225 --PEKIYEKLNAGARVGNDEFFKLRD-----WWTANR-----ETV-WKAITCNAMGN 271
QY 171 S-----CN-----PSYGALSDY--QEMPRAROKKILEEVAVPL----- 202
DB 272 TYFNATCRGERTKGYCRMDQVPTY---FDIVQYLR-----WFEWAEDFCCKKN 321
QY 203 KSVEDIYERFR-----LYC-----LGKLR-----SNFW----- 225
DB 322 KRIDVNRNCRGKDEKDKRYCSRNGYDCEKTKRAIGKRYGKOCISCLYACNRYVDMIN 381
QY 226 SELGLOPETRIINEOLEKINTGFLTINSOPAVNGEKSD---SPYVGSGPGGYVYOKA 282
DB 382 NQKQFPQKQKKYDEIKYTE-----NGASGSRQKRDAGTITTYDQ-----YEKK 429
QY 283 Y-----VEFCSKEKLDALYDKCKDRT--SLTYMAVNDGSKSVNGQTDVNAV 329
DB 430 FYDELINSEKRYVDKFLKLSNEICTKYKDEGGTIDFKVNSD---STSGASGTNVE 485
QY 330 TWGVFPKKEIIOF-----TIYDPSVFNWKRDEAFELMSRGMASLY---PEDEASRKL 379
DB 486 SQGTFYRSKCYQCPGYGVKRVNNGSSNMEWEKKNNGCKSG--KLYPRKDKGTTITTI 543
QY 380 EGVGSHFLVSLVDN---DYINGD 400
DB 544 LKSGGHDIIEKLNKFCDEKNGD 567

RESULT 12
US-08-646-695-2
; Sequence 2, Application US/08646695
; Patent No. 6168943
; GENERAL INFORMATION:
; APPLICANT: Rose, John K.
; APPLICANT: RECOMBINANT VESICULOVIRUSES AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,695
FILING DATE: On Even Date Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-695-2

Query Match 4.4%; Score 97; DB 4; Length 422;
Best Local Similarity 19.7%; Pred. No. 0.22;
Matches 63; Conservative 43; Mismatches 90; Indels 124; Gaps 16;

QY 127 PRRPANEVFKEDYRPIFMANRPS-----YISRTIGMDYPIHGRWDSGNPS----- 175
DB 24 PVEYPADYFRKKEI-PLY-INTTKSLDLRGVYQGL-----KSGNVSIITHVN 70
QY 176 ---YGALSDYQFMPPRRARDKLVEMA-----VPLKSEDIYE----- 210
DB 71 SYLYGALKDIR-----GKLDKDWSSFGINIGKAGDTIGIFDLVSLKALDGLVLDGVS 122
QY 211 -----RFRLYCLGKLRSPWSE--LDGLQPEYKIINEQLEKINTKGFLLT 253
DB 123 DASRTSADDKWLPLYLLGLYRGVGTQMPYRKKLMDGLTNCCKMINQEFPLVPEG---- 178
QY 254 NSQPAVNGEKSDSPYWGSGPGGYVYOKAYVEF-----CSKEKLDALVYCKDRS 305
DB 179 -----RDIFDY-WGNDSNYTKIYAADVMEFHMKKEHCASFYGTIVSRKDCAA 227
QY 306 LTYMAVNNKDSWKSNGOTDVNAVTV-----GVFPAKEI-----IQPTIVD-- 346
DB 228 LATF-----GHLCKITGMSTEDVTWILNREVADENYQMLPGQELDKADSYMPYLLDFG 282
QY 347 ---PVSEFNWKDEAFETMSR 363
DB 283 LSKSPYSYVKNPAPHFHWQ 302

RESULT 13
PCT-US96-06053-2
Sequence 2, Application PC/TUS9606053
GENERAL INFORMATION:
APPLICANT: Yale University
TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
TITLE OF INVENTION: USES
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06053
FILING DATE: 01-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-009-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-06053-2

Query Match 4.4%; Score 97; DB 5; Length 422;
Best Local Similarity 19.7%; Pred. No. 0.22;
Matches 63; Conservative 43; Mismatches 90; Indels 124; Gaps 16;

QY 127 PRRPANEVFKEDYRPIFMANRPS-----YISRTIGMDYPIHGRWDSGNPS----- 175
DB 24 PVEYPADYFRKKEI-PLY-INTTKSLDLRGVYQGL-----KSGNVSIITHVN 70
QY 176 ---YGALSDYQFMPPRRARDKLVEMA-----VPLKSEDIYE----- 210
DB 71 SYLYGALKDIR-----GKLDKDWSSFGINIGKAGDTIGIFDLVSLKALDGLVLDGVS 122
QY 211 -----RFRLYCLGKLRSPWSE--LDGLQPEYKIINEQLEKINTKGFLLT 253
DB 123 DASRTSADDKWLPLYLLGLYRGVGTQMPYRKKLMDGLTNCCKMINQEFPLVPEG---- 178
QY 254 NSQPAVNGEKSDSPYWGSGPGGYVYOKAYVEF-----CSKEKLDALVYCKDRS 305
DB 179 -----RDIFDY-WGNDSNYTKIYAADVMEFHMKKEHCASFYGTIVSRKDCAA 227
QY 306 LTYMAVNNKDSWKSNGOTDVNAVTV-----GVFPAKEI-----IQPTIVD-- 346
DB 228 LATF-----GHLCKITGMSTEDVTWILNREVADENYQMLPGQELDKADSYMPYLLDFG 282
QY 347 ---PVSEFNWKDEAFETMSR 363
DB 283 LSKSPYSYVKNPAPHFHWQ 302

RESULT 14
US-08-965-762-23
Sequence 23, Application US/08965762
Patent No. 6280963
GENERAL INFORMATION:
APPLICANT: Kollip, Vijal
TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
FILE REFERENCE: 0734/062001
CURRENT APPLICATION NUMBER: US/08/965,762
CURRENT FILING DATE: 1997-11-07
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 347
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-08-965-762-23

Query Match 4.2%; Score 91; DB 4; Length 347;
Best Local Similarity 21.3%; Pred. No. 0.65;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 20:30:11 ; Search time 81 Seconds

(without alignments)
1037.867 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187
Sequence: 1 ARVYVQGFYDIDLEKFN.....VELVNDYINGDLFAVEADF 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:.*
2: SP bacteria:.*
3: SP fungi:.*
4: SP human:.*
5: SP invertebrate:.*
6: SP mammal:.*
7: SP mhc:.*
8: SP organelle:.*
9: SP phage:.*
10: SP plant:.*
11: SP rodent:.*
12: SP virus:.*
13: SP vertebrate:.*
14: SP unclassified:.*
15: SP virus:.*
16: SP bacteriophage:.*
17: SP archaea:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 1834 | 83.9 | 594 | 10 Q9SE59 | Q9SE59 arabidopsis |
| 2 | 1834 | 83.9 | 594 | 10 Q94J21 | Q94J21 arabidopsis |
| 3 | 1831 | 83.7 | 594 | 10 Q9SU00 | Q9SU00 arabidopsis |
| 4 | 1777 | 81.3 | 592 | 10 Q9SE60 | Q9SE60 arabidopsis |
| 5 | 1762 | 80.6 | 593 | 10 Q9SE94 | Q9SE94 zea mays (m |
| 6 | 246 | 11.2 | 276 | 16 Q8Y389 | Q8Y389 ralsdonia s |
| 7 | 222 | 10.2 | 290 | 16 Q91687 | Q91687 pseudomonas |
| 8 | 221 | 10.1 | 314 | 16 Q9A6F4 | Q9A6F4 caudobacter |
| 9 | 215 | 9.8 | 275 | 16 Q9PEA7 | Q9PEA7 xylella fas |
| 10 | 198 | 9.1 | 313 | 16 Q92NKL | Q92NKL rhizobium m |
| 11 | 197.5 | 9.0 | 163 | 3 Q94090 | Q94090 saccharomyc |
| 12 | 196 | 9.0 | 295 | 16 Q9CP31 | Q9CP31 pasteurella |
| 13 | 193 | 8.8 | 307 | 16 Q9S2V3 | Q9S2V3 streptomyces |
| 14 | 189 | 8.6 | 296 | 16 Q92302 | Q92302 salmoneilla |
| 15 | 189 | 8.6 | 296 | 16 Q8X766 | Q8X766 escherichia |
| 16 | 188 | 8.6 | 282 | 16 Q9PN93 | Q9PN93 campylobact |

| | | | | | |
|----|-------|-----|------|-----------|---------------------|
| 17 | 186 | 8.5 | 293 | 2 Q9LSC1 | Q9LSC1 leptospira |
| 18 | 184 | 8.4 | 183 | 6 Q9NDY5 | Q9NDY5 sus acrofa |
| 19 | 182 | 8.3 | 292 | 16 Q9JZ02 | Q9JZ02 neisseria m |
| 20 | 182 | 8.3 | 292 | 16 Q9JZ07 | Q9JZ07 neisseria m |
| 21 | 182 | 8.3 | 292 | 2 Q93ER8 | Q93ER8 photobacter |
| 22 | 179 | 8.2 | 314 | 16 Q8ZJ16 | Q8ZJ16 yersinia pe |
| 23 | 171 | 7.8 | 304 | 16 Q9KNP6 | Q9KNP6 agrobacteri |
| 24 | 163 | 7.5 | 283 | 16 Q9CG56 | Q9CG56 vibrio chol |
| 25 | 162.5 | 7.4 | 305 | 16 Q9K8K7 | Q9K8K7 lactococcus |
| 26 | 162 | 7.4 | 274 | 16 Q9K186 | Q9K186 rhizobium l |
| 27 | 155.5 | 7.1 | 349 | 5 Q9V710 | Q9V710 bruceella me |
| 28 | 154.5 | 6.8 | 130 | 2 Q9ACM5 | Q9ACM5 streptococ |
| 29 | 148.5 | 6.8 | 288 | 16 Q9T530 | Q9T530 streptococ |
| 30 | 142 | 6.3 | 318 | 2 Q9F354 | Q9F354 rhodotherm |
| 31 | 138 | 6.3 | 305 | 2 Q938W5 | Q938W5 streptomyc |
| 32 | 138 | 6.3 | 674 | 17 Q8TJ17 | Q8TJ17 methanosarc |
| 33 | 108 | 4.8 | 1138 | 16 Q9PK07 | Q9PK07 chlamydia m |
| 34 | 104.5 | 4.8 | 483 | 12 Q9ODJ3 | Q9ODJ3 clover yell |
| 35 | 104 | 4.8 | 538 | 12 Q67458 | Q67458 figwort mos |
| 36 | 103 | 4.7 | 706 | 5 Q9NDE9 | Q9NDE9 dictyostell |
| 37 | 102 | 4.7 | 920 | 13 Q910A4 | Q910A4 brachydanio |
| 38 | 102 | 4.7 | 301 | 16 Q8R926 | Q8R926 thermotanaer |
| 39 | 101 | 4.6 | 548 | 3 Q9P8J5 | Q9P8J5 candida alb |
| 40 | 101 | 4.6 | 422 | 12 Q88993 | Q88993 vesicular s |
| 41 | 100 | 4.6 | 1142 | 16 Q97J24 | Q97J24 clostridium |
| 42 | 100 | 4.5 | 619 | 5 Q9VMI9 | Q9VMI9 dirosophila |
| 43 | 99.5 | 4.5 | 406 | 12 Q83136 | Q83136 bean yellow |
| 44 | 99 | 4.5 | 582 | 16 Q67141 | Q67141 aquilex aeo |
| 45 | 98.5 | 4.5 | | | |

ALIGNMENTS

RESULT 1
Q9SE59 PRELIMINARY: PRT; 594 AA.
ID Q9SE59;
AC Q9SE59;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Methyltetrahydrofolate reductase MTHFR2 (EC 1.5.1.20)
DE (At2g44160/F6EL3.29).
DE Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20062814; PubMed=10593891;
RA Koje S., Wang H., McNeil S.D., Raymond R.K., Appling D.R.,
RA Snachar-Hill Y., Bohner H.J., Hanson A.D.;
RA "Isolation, characterization and functional expression of cDNAs
RT encoding NADH-dependent methyltetrahydrofolate reductase from
RT higher plants.";
RL J. Biol. Chem. 274:36089-36096(1999).
RN [2]
RN SEQUENCE FROM N.A.
RA Shih P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Ouedera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi W., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF181967; AAD5788.1; -
DR EMBL; AY050434; AAK91450.1; -
DR HSSP; P00394; 1B5T.
DR InterPro; IPR000194; ATPase_a/bcentre.

```

DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003171; Mchydrof_redctse.
DR Pfam: PF02219; MTHFR_1.
DR TIGRfams: TIGR00677; fadh2_euk; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR OXfoldreductase.
SQ SEQUENCE 594 AA: 66802 MW: E2FAE919CE10E98 CRC64;

Query Match      83.9%: Score 1834; DB 10; Length 594;
Best Local Similarity 82.1%: Pred. No. 1.6e-148;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

OY 1 ARVYVOLFVDFDIFLKFVNDRCRQIGITCPYVGPIMPINNYKGFIRMTGFCRTKIPADIM 60
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 187 ADLVITOLFVDFDIFLKFVNDRCRQIGISCPYVGPIMPINNYKGFIRMTGFCRTKIPVEVM 246
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 61 AALEPIKDNBEAVKAYGIHLGTECKKILAHGIKTLHLTYLNMKSKALATLNMGLIEES 120
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 247 AALEPIKDNBEAVKAYGIHLGTECKKMLAHGKSLHLTYLNMKSKALATLNMGLIMDES 306
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 121 KVSRLSPWRRPANVFRVKEVDRPIFMANRPKSYISRTIGMDQYHGWGSCNPSYGALS 180
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 307 KISRLSPWRRPANVFRKEDVRPIFMANRPKSYISRTIGMDPFGQWGDSSASAYGALS 366
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 181 DYOFMPRAROKKILQEMAVPLKSVEDIYERPRILGCKLNSNWSLDELQEPETRIINE 240
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 367 DHQFSRRAROKKILQDEMVAFLKSVEDIQERKELCLGNLKSFWSELDELQEPETRIINE 426
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 241 QLEKITKGFLLTINSOPAVNGEKSDSPYWGSGPGGYVYQKAYVEFCSEKRLDALVDKC 300
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 427 QLIKVMSKGFLLTINSOPVNAERSDSPYWGSGPGGYVYQKAYVEFCSEKRLDAVVEKC 486
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 301 KDRTSLTYMAVNVKDGSMKSNVGTDVNAVTWGVFPKKEIIQPTITVDPSFNWVKDEAFET 360
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 487 KALPSTTYMAVNVKGEQVNSNTAQADVAATWGVFPKKEIIQPTITVDPSFNWVKDEAFET 546
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 361 WSRGNASLYPDEDEASRKLYEEVGSHFLVSLVNDNYINGDLEFAFVAD 407
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 547 WSRSMANLYPEADPSRNILLEVKNSYVSLVLENDYINGDLEFAFVAD 593
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

RESULT 2
O94JZ1 PRELIMINARY; PRT; 594 AA.
AC O94JZ1:
DC 01-DEC-2001 (TREMblrel_19, Created)
DT 01-DEC-2001 (TREMblrel_19, Last sequence update)
DT 01-JUN-2002 (TREMblrel_21, Last annotation update)
DE Putative methyleneretrahydrofolate reductase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bann J., Garinaci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kania A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sekano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
RA Ecker J., Theologis A., Davis R.W.,
RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ370515; AAK3892.1; -.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003171; Mchydrof_redctse.
DR Pfam: PF02219; MTHFR_1.
DR TIGRfams: TIGR00677; fadh2_euk; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR OXfoldreductase.
SQ SEQUENCE 594 AA: 66842 MW: 14D664CFDE7CD8DE CRC64;

Query Match      83.9%: Score 1834; DB 10; Length 594;

```

```

Best Local Similarity 82.1%: Pred. No. 1.6e-148;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

OY 1 ARVYVOLFVDFDIFLKFVNDRCRQIGITCPYVGPIMPINNYKGFIRMTGFCRTKIPADIM 60
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 187 ADLVITOLFVDFDIFLKFVNDRCRQIGISCPYVGPIMPINNYKGFIRMTGFCRTKIPVEVM 246
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 61 AALEPIKDNBEAVKAYGIHLGTECKKILAHGIKTLHLTYLNMKSKALATLNMGLIEES 120
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 247 AALEPIKDNBEAVKAYGIHLGTECKKMLAHGKSLHLTYLNMKSKALATLNMGLIMDES 306
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 121 KVSRLSPWRRPANVFRVKEVDRPIFMANRPKSYISRTIGMDQYHGWGSCNPSYGALS 180
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 307 KISRLSPWRRPANVFRKEDVRPIFMANRPKSYISRTIGMDPFGQWGDSSASAYGALS 366
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 181 DYOFMPRAROKKILQEMAVPLKSVEDIYERPRILGCKLNSNWSLDELQEPETRIINE 240
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 367 DHQFSRRAROKKILQDEMVAFLKSVEDIQERKELCLGNLKSFWSELDELQEPETRIINE 426
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 241 QLEKITKGFLLTINSOPAVNGEKSDSPYWGSGPGGYVYQKAYVEFCSEKRLDALVDKC 300
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 427 QLIKVMSKGFLLTINSOPVNAERSDSPYWGSGPGGYVYQKAYVEFCSEKRLDAVVEKC 486
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 301 KDRTSLTYMAVNVKDGSMKSNVGTDVNAVTWGVFPKKEIIQPTITVDPSFNWVKDEAFET 360
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 487 KALPSTTYMAVNVKGEQVNSNTAQADVAATWGVFPKKEIIQPTITVDPSFNWVKDEAFET 546
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 361 WSRGNASLYPDEDEASRKLYEEVGSHFLVSLVNDNYINGDLEFAFVAD 407
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 547 WSRSMANLYPEADPSRNILLEVKNSYVSLVLENDYINGDLEFAFVAD 593
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

RESULT 3
O9SUJ0 PRELIMINARY; PRT; 594 AA.
AC O9SUJ0:
DC 01-MAY-2000 (TREMblrel_13, Created)
DT 01-MAY-2000 (TREMblrel_13, Last sequence update)
DT 01-JUN-2002 (TREMblrel_21, Last annotation update)
DE Methyleneretrahydrofolate reductase (Ec 1.5.1.20).
GN MTHFR1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Ravanel S., Rebelle F., Douce R.;
RT "Folate metabolism in higher plants: cloning of a cDNA for 5,10-
RT methylenetetrahydrofolate reductase in Arabidopsis thaliana.";
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ245414; CAB53783.1; -.
DR HESP: P00394; 1B5T.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003171; Mchydrof_redctse.
DR Pfam: PF02219; MTHFR_1.
DR TIGRfams: TIGR00677; fadh2_euk; 1.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR OXfoldreductase.
SQ SEQUENCE 594 AA: 66803 MW: 64859A988DA81657 CRC64;

Query Match      83.7%: Score 1831; DB 10; Length 594;
Best Local Similarity 82.1%: Pred. No. 2.9e-148;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

OY 1 ARVYVOLFVDFDIFLKFVNDRCRQIGITCPYVGPIMPINNYKGFIRMTGFCRTKIPADIM 60
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :
DB 187 ADLVITOLFVDFDIFLKFVNDRCRQIGISCPYVGPIMPINNYKGFIRMTGFCRTKIPVEVM 246
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

OY 61 AALEPIKDNBEAVKAYGIHLGTECKKILAHGIKTLHLTYLNMKSKALATLNMGLIEES 120
   | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : :

```

| | | | |
|----|-----|---|------|
| Db | 247 | AABEIKONEAAVKAAGIHLEGTWCKKMLAHVKSJHLHYLIMBKSAJLIMNLGIMDES | 3066 |
| Qy | 121 | KYSRLPPRRANVPFKEDVAPLEFMANRPKSYSTSTGMOYHGMKSDONSPRYALS | 1808 |
| Db | 307 | KISRLPPRRANVPFKEDVAPLEFMANRPKSYSTSTGKWDPPQGMDSASATGALS | 3666 |
| Qy | 181 | DYQMRPAADKYLEMAVPLKSYVEDIYRFLYCLKLSNWSJLSDLOPEKRIINE | 2400 |
| Db | 367 | DYQSRPAPARKLQDQWVPLKSYVEDIYRFLYCLKLSNWSJLSDLOPEKRIINE | 4266 |
| Qy | 241 | QLEKINTYGLTJNSOPAVNGEKSDSPVWGPGGTYQKAYVEFFCSKEKJLDVDC | 3000 |
| Db | 427 | QLEKINSGLTJNSOPAVNKRSDSPVWGPGGTYQKAYVEFFCSKEKJLDVDC | 4866 |
| Qy | 301 | KDQTSLYMAVNGKDSWKSNGQTDVNAVTWGVFPKKEIOPITLVDFSVSNWKADEFI | 3600 |
| Db | 487 | KALPITLYMAVNGEDQWVNTQADVNAVTWGVFPKKEIOPITLVDFSVSNWKADEET | 5466 |
| Qy | 361 | WSRGNASTLYPDEASRKLVEEYGVSHPLVSLVNDYINGDLEFAVAD | 407 |
| Db | 547 | WSRGNASTLYPDEASRKLVEEYKNSYVSLVNDYINGDLEFAVAD | 593 |

RESULT 4

| ID | OSSE60 | PRELIMINARY: | PRT: | 592 AA. |
|----|--|--------------|------|---------|
| DT | 01-MAY-2000 (TREMblrel. 13, Created) | | | |
| DT | 01-MAY-2000 (TREMblrel. 13, Last sequence update) | | | |
| DT | 01-JUN-2002 (TREMblrel. 21, Last annotation update) | | | |
| DE | Methylmenettrahydrofolate reductase MNHFP1 (EC 1.5.1.20) [Putative methylmenettrahydrofolate reductase MNHFP1]. | | | |
| GN | P24G15.240 OR A35C59970. | | | |
| OS | Arabidopsis thaliana (Mouse-ear cress). | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucrosid II; Brassicales; Brassicaceae; Arabidopsids. | | | |
| OK | NCBI_TaxId=3702; | | | |
| RM | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MedLine-20062814; PubMed-10593991; | | | |
| RA | Rejde S., Wang H., McNeil S.D., Raymond R.K., Appling D.R., Shachar-Hill Y., Boehrert H.J., Hanson A.D.; | | | |
| RA | "Isolation, characterization and functional expression of cDNAs encoding NADH-dependent methylmenettrahydrofolate reductase from higher plants."; | | | |
| RT | | | | |
| RL | J. Biol. Chem. 274:36089-36096(1999). | | | |
| RM | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W., Lemcke K., Meyer R.F.X., Quettier F., Salenoubat M.; | | | |
| RL | Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases. | | | |
| RM | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | EU Arabidopsis sequencing project; | | | |
| RL | Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases. | | | |
| RM | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Yamada K., Banb J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D., Lee J.M., Ondera C.S., Quach H.L., Tang C., Forlun M., Wu H.C., Yamamura T., Yu G., Yu S., Bowser L., Cahninci P., Chen H., Cheuk R., Hayashizaki T., Ishida J., Jones T., Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shimizu P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.; | | | |
| RT | "Full length cDNA of gene At3g59970 (GI:15322215)." | | | |
| RL | Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases. | | | |
| DR | EMBL: AF181966; AAD55787.1; - | | | |
| DR | EMBL: AL186447; CAB75816.1; - | | | |
| DR | EMBL: AY070034; AAL49791.1; - | | | |
| DR | HSSP: P00394; 1B5T. | | | |
| DR | InterPro: IPR000194; ATRase_a/bcentre. | | | |

DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003771; MchHydrof_redctase.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRPFAMS: TIGR00677; Fadh2_euk. 1.
DR PROSITE: PS00152; ATPase_ALPHA_BETA; UNKNOWN. 1.
KW Oxidoreductase.
MO SEQUENCE 592 AA; 66288 MW; 72D7453AF1AF1573 CRC64;

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 81.3%; | Score 1777; | DB 10; | Length 592; |
| Best Local Similarity | 78.6%; | Pred. No. 1.2e-143; | | |
| Matches 319; | Conservative 44; | Mismatches 43; | Indels 0; | Gaps 0; |

```

OY 1 ARVAVOLATPDDIFELKRVNDRCIOIGTCIYVNGIMEINNNKGFIRBMTCGCKIRPADIM 60
Db 187 ADLVLVOLFDDIDFLKRVNDCROIGNCEIYVIMIPISNNKGFIRBMAGFCIKIRIAELT 246
OY 61 AALEPIKDEEAVKAVGIHLJOTEMCKILAHGKITLHLTYLTNNKXSALAILMNLGLEBS 120
Db 247 AALEPIKDNDEAVKAVGIHFATEMCKILAHGTTLSHLTYLTNVKDSAGILMNLGLIDES 306
OY 121 KVSLSLPMKRPANVRYVEDVRFPMANRKSYSITSTICMOOYPHGRMGSCNDPSGALS 180
Db 307 KISHSILPMKRPANVRETRVEDVRFPMANRKSYSITSTIKCMANDFPHGRMGDSHSAISITS 366
OY 181 DYOPMRPRADKKLVEEAVPLKSVEDIYERFLVCLGKRSNPSWSEIDJLOPETKIIME 240
Db 367 DYOPMRPRADKKLQOEAVVPLKSIIEYDERFELCIGMLKSSPSEIDJLOPETKIIME 426
OY 241 QLEKINTKFLVINSOPAVNNEKSDSTPYWGMGPFGVYQKAVYEFFECSREKILDAVLNDC 300
Db 427 OLGRINSNGFLTINSOPVNAKAGDSPALIGMGPFGVYQKAYLEEFCSKODLTILEKS 486
OY 301 KDRSTLTLYAAVNNKDSNMKSNNGOVDVAAYWGYEPAKKEILOTYTIDVPSFNWMDAEFEI 360
Db 487 KAPPSITLYAAVNNKSEWNSNNGCEDVAAYWGYEPAKEVIOPTYTIDVPSFRYWKDAFEI 546
OY 361 WSRGASLVPDEDEASRKLVEEVSSEHLVSLVNDYDINGDLFAVFA 406
Db 547 WSRGASLVPDEDEASRKLLEEVKSNYSIVSLVNDYDINDIDISVFA 592

```

RESULT 5

Q9SE94 ID Q9SE94 PRELIMINARY; 593 AA.
AC Q9SE94;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Methylentetrahydrofolate reductase (EC 1.5.1.20).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20062814; PubMed=1053891;
RA Roje S., Mang H., McNeill S.D., Raymond R.K., Appling D.R.,
RA Shachar-Hill Y., Bohmelt H.J., Hanson A.D.;
RT Isolation, characterization and functional expression of cDNAs
RT encoding NADH-dependent methylentetrahydrofolate reductase from
RT higher plants.";
RL J. Biol. Chem. 274:36089-36096(1999).
DR EMBL: AF174486; AAD51733.1; -.
DR HSSP: P00394; 1B5T.
DR InterPro: IPR000194; ATPase_a/bcentre.
DR InterPro: IPR004621; Pcdh2_euk.
DR InterPro: IPR003171; Methylotf_redcse.
DR Pfam: PF02219; MYERF_1.
DR TIGRfams: TIGR00677; fcdh2_euk; 1.
DR PROSITE: PS00152; ATPase_ALPHA_BETA; UNKNOWN_1.
SQ Oxidoreductase. KW
SEQUENCE 593 AA; 66429 MW; 456899675B638B4 CRC64;

RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Elsen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeJoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Outback T., Tran K., Wolf A., Vamthavan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
 RT "Complete genome sequence of *Caulobacter crescentus*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005886; AAK24111.1; -.
 DR HSSP: P00394; 1B5T.
 DR TIGR: CC2140; -.
 DR InterPro: IPR004620; Fadh2_bact.
 DR InterPro: IPR003171; Methylrof_redctse.
 DR Pfam: PF02219; MTHFR; 1.
 DR TIGRFAMS: TIGR00676; fadh2; 1.
 KM Complete proteome.
 SQ SEQUENCE 314 AA; 34247 MW; 521E7A352E34A9C5 CRC64;

Query Match 10.1%; Score 221; DB 16; Length 314;
 Best Local Similarity 37.9%; Pred. No. 1,le-10;
 Matches 44; Conservative 21; Mismatches 51; Indels 0; Gaps 0;

OY 5 VTQLFYDTDFLEKFNVDRCQIGTCIPYVGIMPINNYKGFTRMGFCKTKIPADIMALE 64
 DB 191 ISOEFFDADLRFDYDKVRAGITIPYGPIMPVTNFAGLKKMAACQTAIPSWGLNLF 250
 OY 65 PIKDNBEAVKAYGIHLGTEMCKKILAHGKIKTLHLTYLNMESALAIIMNGLIEES 120
 DB 251 GIENDAETRRILACSVAAEMCAKLOEGEDFPHYTLNRADLVYALICRVGVRIS 306

RESULT 9

O9PEA7 PRELIMINARY; PRT: 275 AA.
 AC O9PEA7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE 5,10-methylenetetrahydrofolate reductase.
 GN Xf1121.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carriero D.M., Carrier H.,
 RA Colaudo N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gubler A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Macneil M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Norega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,

RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Melandis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 DR EMBL: AE003948; AAF83931.1; -.
 DR HSSP: P00394; 1B5T.
 DR InterPro: IPR003171; Methylrof_redctse.
 DR Pfam: PF02219; MTHFR; 1.
 KM Complete proteome.
 SQ SEQUENCE 275 AA; 30499 MW; 22D6C156B53241D0 CRC64;

Query Match 9.8%; Score 215; DB 16; Length 275;
 Best Local Similarity 33.9%; Pred. No. 2,9e-10;
 Matches 39; Conservative 30; Mismatches 46; Indels 0; Gaps 0;

OY 1 ARVYVQLFYDTDFLEKFNVDRCQIGTCIPYVGIMPINNYKGFTRMGFCKTKIPADIM 60
 DB 160 ADAATQYFNPDPAYPHFVDAVGRIGVTIPYAGVPIPNFRLHHSQCGAEIPRWIT 219
 OY 61 AALEPIKDNBEAVKAYGIHLGTEMCKKILAHGKIKTLHLTYLNMESALAIIMNG 115
 DB 220 KKNQAVGDDIKSIRAFGADVYVALCERLLAGAPGLHFTYTLNARFSTGYLQRLG 274

RESULT 10

O92NKL PRELIMINARY; PRT: 313 AA.

AC O92NKL;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Probable 5,10-methylenetetrahydrofolate reductase oxidoreductase
 DE protein (EC 1.7.99.5).
 GN METE OR R02199 OR SMC01843.
 GN Rhizobium meliloti (Sinorhizobium meliloti).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsberger U.,
 RA Renard C., Thebaud P., Vandendol M., Weidner S., Gallibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 Sinorhizobium meliloti strain 1021.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 RL EMBL: AL591789; CAC46778.1; -.
 DR InterPro: IPR004620; Fadh2_bact.
 DR InterPro: IPR003171; Methylrof_redctse.
 DR Pfam: PF02219; MTHFR; 1.
 DR TIGRFAMS: TIGR00676; fadh2; 1.
 KM Oxidoreductase; Complete proteome.
 SQ SEQUENCE 313 AA; 34196 MW; C8BP3136B0D15071 CRC64;

Query Match 9.1%; Score 198; DB 16; Length 313;
 Best Local Similarity 34.2%; Pred. No. 1e-08;
 Matches 38; Conservative 20; Mismatches 53; Indels 0; Gaps 0;

OY 5 VTQLFYDTDFLEKFNVDRCQIGTCIPYVGIMPINNYKGFTRMGFCKTKIPADIMALE 64
 DB 180 LTOFFPNDFFERYLERYVRAGIAITIVPEIILINMLUYQKFAKGCARVESPVSRLG 239
 OY 65 PIKDNBEAVKAYGIHLGTEMCKKILAHGKIKTLHLTYLNMESALAIIMNG 115
 DB 240 PIDDQPEEREREAHIAAEQVADLVRRGVSDFLTYMNRQILITAVCEIMG 290

RESULT 11

O94090

| | | | | |
|----|--|--------------|-----------|-------------------------|
| ID | 094090. | PRELIMINARY; | PRT; | 163 AA. |
| AC | 094090. | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Created) | | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Last sequence update) | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) | | | |
| DE | Terra hydrofolate reductase (fragment). | | | |
| GN | MTKL. | | | |
| OS | Saccharomyces cerevisiae (Baker's yeast). | | | |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | | | |
| CC | Saccharomycetales; Saccharomycetaceae; Saccharomycetes. | | | |
| OX | NCBI_TaxID=4932; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 28383; | | | |
| RA | Housen I., Lafontaine D., Belot N., Vandenhautte J.; | | | |
| RL | Submitted (APR-1995) to the EMBL/Genbank/DBJ databases. | | | |
| DR | EMBL; U24271; AAC9805.1; -. | | | |
| DR | HSSP; P00394; 1B5T. | | | |
| DR | InterPro; IPR003171; Mchydrof__redctse. | | | |
| DR | Pfam; PF02219; MTHFR.1. | | | |
| FT | NON_TER | 1 | 1 | |
| ET | NON_TER | 163 | 163 | |
| SO | SEQUENCE | 163 AA; | 18920 MW; | 58412E4B893D069C CRC64; |

| | | | | | |
|----|-----------------------|---|--------------------|-------|----------------------------------|
| | Query Match | 9.0% | Score 197.5; | DB 3; | length 163; |
| | Best Local Similarity | 39.6%; | Pred. No. 4.3e-09; | | |
| | Matches | 40; | Conservative | 22; | Mismatches 36; Indels 3; Gaps 3. |
| OY | 4 | IYTDLFYPTDIFLAKVNDRCRGITGCTPIYPGIMPINNKGFTIRMGECKTKIPADIMAL | 63 | | |
| | | ::::: ::: | : | : | : |
| Db | 65 | ITTFHFVDNEINNNCSOVRAGMDLPITPGIMPITTYAAFLRRQWAMLNH-PSTFVLV | 123 | | |
| | | ::::: ::: | : | : | : |
| OY | 64 | EPIKNDIEAVKAYGTIHLSTEMCKKLLAIGIKT-LHLVYLNM | 103 | | |
| | | ::::: ::: | : | : | : |
| Db | 124 | DPKIDDELIVRDIGTNILIVE-CVKRAQRWLRSRLHIITYNNL | 163 | | |
| | | ::::: ::: | : | : | : |

| | | | |
|-----------|---|--------------|--------------|
| RESULT 12 | | | |
| Q9CP31 | ID | PRELIMINARY; | PRT; 295 AA. |
| Q9CP31 | Q9CP31 | | |
| AC | Q9CP31; | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Created) | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) | | |
| DE | MefF. | | |
| GN | MEFF OR PM0235. | | |
| OS | Pasteurella multocida. | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; | | |
| OC | Pasteurella. | | |
| OX | NCBI_TaxID=747; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=PM70; | | |
| RX | MEDLINE=21145866; PubMed=11248100; | | |
| RA | May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.; | | |
| RT | "Complete genomic sequence of Pasteurella multocida pm70."; | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001). | | |
| DR | EMBL; AE006058; AAK02319.1; . | | |
| DR | HSPD; P00394; IBSY. | | |
| DR | InterPro; IPR004620; Fadh2_bact. | | |
| DR | InterPro; IPR003171; Methylrof_redctse. | | |
| DR | Pfam; PF00219; MTHFR; 1. | | |
| DR | TIGRFAMS; TIGR00676; fadh2; 1. | | |
| Q0 | Complete proteome. | | |
| Q0 | SEQUENCE 295 AA; 33064 MW; 7BC42A982E03AFEA CRC64; | | |

| | | | | | |
|----|--|------------------|--------------------|-----------|-------------|
| | Query Match | 9.0% | Score 196: | DB 16: | Length 295; |
| | Best Local Similarity | 32.88% | Pred. No. 1.4e-08; | | |
| Ob | Matches 38: | Conservative 21: | Mismatches 57: | Indels 0; | Gaps 0; |
| Oy | 1 ARVVTQLFVDYDPLFLKRVNDCRQIGTICPTVPGLMNNYKGIKRMGTGCKTKIPADIM 60 | | | | |
| | - - - : : : : - - - - - : : : : : - - - - - : : : : : | | | | |
| Ob | 175 ANHHVITOFFEFDISLRTDRRCASIGIDAEIVPGILPTNRNKAKMAATNVKKIPAVIA 234 | | | | |
| | - - - - - : : : : - - - - - : : : : - - - - - : : : : - - - - - | | | | |

Qy 61 AALEPIKDNEAAVKA YGIHGTETMCKRILANGIKTLHLVLTLMNEXSALAILNLGL 116
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 235 KNYEGLDDEPTTRNLVAASVAMDNAVYLSREGVNDHFHTLNRSELTYYAICHTTIGI 290

RESULT 13

ID Q9S2V3 PRELIMINARY; PRT; 307 AA.

DT 01-MAY-2000 (TREMBLRel. 13, Created)

DT 01-JUN-2002 (TREMBLE1. 21, Last annotation update)

22 5/10 mcmf2ccccccfjao1010acc 1ccaccacc;
GN METF OR SCO2103 OR SC4A10.36C.

Bacteria; Firmicutes; Actinobacteria

NCBI_TaxID=1902; streptomycetaceae, streptomycetes

| RP | SEQUENCE FROM N.A. |
|----|--------------------|
| RM | [11] |

| | | |
|----|---------------|-----|
| RC | STRALE=A3(2); | |
| BA | Saunders D C | HAY |

Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
PC SMPAIN-23(2).

RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (AUG-1000) to the EMBL/Genbank/DBIT databases

RN [3]

```
RC STRAIN=A3(2);
```

RA Redenbach M., Kies

RT "A set of ordered cosmids and a detailed genetic and physical map for

RL Mol. Microbiol. 21

RP SEQUENCE FROM N.A.

RA Bentley S.D., Chat

RA Harper D., Bateman

RA Huang C.-H., Kiese

Seeger K., Saunderson

RA Hopwood D.A.;

| RT | complete genome | coellicolor A3(2). " |
|----|-----------------|----------------------|
| R1 | | |

KL Nature 417:141-147
DR EMBL: AL109663: CA

```
DR HSSP; P00394; LB51
DR InterPro: IPR00462
```

DR InterPro: IPR00317
DR Pfam: PF02219. MTH

DR TIGREAMS; TIGR0067
SC SEOTENCE 307 AA

QUESTIONS

Best Local Similarity

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

—
—
• •
—
—
• •
• •
—

Q C O T C A S E F C C


```

RESULT 14
082302
ID 082302 PRELIMINARY; PRT; 296 AA.
AC 082302;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase.
GN str3761.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Holden M.T.G., Sebahia M.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Kirogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627279; CADD09516.1; -
DR InterPro; IPR004620; Fadh2_bact.
DR InterPro; IPR003171; Methylrof_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00676; fadh2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33202 MW; 3165CE17678ACBBE CRC64;

Query Match 8.6%; Score 189; DB 16; Length 296;
Best Local Similarity 32.1%; Pred. No. 5.4e-08;
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;

QY 5 VTQLEFYDTDFLEKFNVDRCQIGITCPVIGIMPINNYKGFIRMTGFCCKTIPADIMALE 64
DB 181 ITQFFEDVESYLRFRDRCVSAGIDVELIPGILPVSFKQAKKFPADMTNVRIRPAMMAQMD 240

QY 65 PIRKNEAVKAVGIHLGTECKKTIAGIKTLHLTYTLNMEKSALAILMNLGL 116
DB 241 GIDDDAETRKLVGANIAMDVKILISREGVDFEFTLNRAEMSYAICHTLGV 292

RESULT 15
08X766
ID 08X766 PRELIMINARY; PRT; 296 AA.
AC 08X766;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase.
GN MTFP OR Z5496 OR ECS4870.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

```

```

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005625; AAG59142.1; -
DR InterPro; IPR002567; BAB38293.1; -
DR InterPro; IPR004620; Fadh2_bact.
DR InterPro; IPR003171; Methylrof_redctse.
DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00676; fadh2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33129 MW; 1FA66221D8296676 CRC64;

Query Match 8.6%; Score 189; DB 16; Length 296;
Best Local Similarity 32.1%; Pred. No. 5.4e-08;
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;

QY 5 VTQLEFYDTDFLEKFNVDRCQIGITCPVIGIMPINNYKGFIRMTGFCCKTIPADIMALE 64
DB 181 ITQFFEDVESYLRFRDRCVSAGIDVELIPGILPVSFKQAKKFPADMTNVRIRPAMMAQMD 240

QY 65 PIRKNEAVKAVGIHLGTECKKTIAGIKTLHLTYTLNMEKSALAILMNLGL 116
DB 241 GIDDDAETRKLVGANIAMDVKILISREGVDFEFTLNRAEMSYAICHTLGV 292

```

Search completed: February 14, 2003, 23:09:55
Job time : 86 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: February 14, 2003, 23:08:27 ; Search time 2159 Seconds
(without alignments)
3060.564 Million cell updates/sec

Title: US-09-720-451-6
Perfect score: 2187
Sequence: 1 ARVLTQLEFYDTDFLEKFN.....VSLVDNDYINGDLFAVFAADF 408

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=x1h
-O=/cgn2.1/USPTO_SPOOL/US09720451/runat_05022003_074816_2604/app.query.fasta.1.583
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-OCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09720451_@cgn.1.1_761_@runat_05022003_074816_2604 -NCPU=6 -ICPU=3
-NO_XLUPX -NO_MMAP -LARGEQUERY -NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_estbhum:*
3: em_estlin:*
4: em_esttmu:*
5: em_estlov:*
6: em_estopl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|----------------------|
| 1 | 1750 | 80.0 | 2423 | 11 | AY109096 | AY109096 Zea mays |
| 2 | 1400 | 64.0 | 851 | 14 | BQ165347 | BQ165347 EST611216 |
| 3 | 1216 | 55.6 | 759 | 13 | B1308605 | B1308605 EST530015 |
| 4 | 1129 | 51.6 | 934 | 12 | BG365861 | BG365861 HVSME1000 |
| 5 | 1128 | 51.6 | 823 | 12 | BC418616 | BC418616 HVSMEK002 |
| 6 | 1104 | 50.5 | 666 | 10 | AW201065 | AW201065 seq7d12.y |
| 7 | 1097 | 50.2 | 760 | 14 | BO994876 | BO994876 GGF8G21.y |
| 8 | 1093 | 50.0 | 621 | 10 | AM830074 | AM830074 sm22h02.y |
| 9 | 1087 | 49.7 | 700 | 10 | BE576981 | BE576981 LO-1586T3 |
| 10 | 1079.5 | 49.4 | 752 | 13 | BI421654 | BI421654 EST532320 |
| 11 | 1052.5 | 48.1 | 842 | 12 | BC418214 | BC418214 HVSMEK002 |
| 12 | 1052 | 48.1 | 611 | 12 | BE805904 | BE805904 ss62h07.y |
| 13 | 1049 | 48.0 | 729 | 12 | BC123885 | BC123885 EST469531 |
| 14 | 1045 | 47.8 | 821 | 10 | BE643238 | BE643238 Cr12_8_E1 |
| 15 | 1042 | 47.6 | 588 | 12 | BF071278 | BF071278 st42e03.y |
| 16 | 1029 | 47.1 | 755 | 13 | BI308302 | BI308302 EST529712 |
| 17 | 1021 | 46.7 | 911 | 12 | BE231181 | BE231181 HVSME9001 |
| 18 | 1017.5 | 46.5 | 836 | 12 | BC418185 | BC418185 HVSMEK002 |
| 19 | 1012 | 46.3 | 573 | 10 | AW203249 | AW203249 sf27g11.y |
| 20 | 1011 | 46.2 | 579 | 10 | AW203207 | AW203207 sf27d02.y |
| 21 | 1011 | 46.2 | 652 | 12 | BG904512 | BG904512 Talr1132g |
| 22 | 1008 | 46.1 | 668 | 14 | BQ404086 | BQ404086 GA_Ed006 |
| 23 | 1004 | 45.9 | 568 | 10 | BE346411 | BE346411 sp25c05.y |
| 24 | 996 | 45.5 | 573 | 10 | AW703738 | AW703738 sk23g05.y |
| 25 | 995 | 45.5 | 569 | 14 | BM885389 | BM885389 sal19b011.y |
| 26 | 980 | 44.8 | 587 | 10 | AW61036 | AW61036 sl17b07.y |
| 27 | 980 | 44.8 | 593 | 10 | BE204232 | BE204232 EST396908 |
| 28 | 976 | 44.6 | 627 | 10 | AV915401 | AV915401 AV915401 |
| 29 | 966 | 44.2 | 557 | 13 | BI425670 | BI425670 sal7b0b07.y |
| 30 | 965 | 44.1 | 634 | 10 | BE346389 | BE346389 sp25a05.y |
| 31 | 960 | 43.9 | 575 | 10 | AW733729 | AW733729 sk77e01.y |
| 32 | 955 | 43.7 | 757 | 17 | BH012972 | BH012972 TDGAQ75TH |
| 33 | 947 | 43.3 | 676 | 10 | BE403890 | BE403890 WHE0420_H |
| 34 | 944 | 43.2 | 565 | 9 | AI190049 | AI190049 sc07a01.y |
| 35 | 943.5 | 43.1 | 721 | 14 | BO993154 | BO993154 OGF2a01.y |
| 36 | 936 | 42.8 | 550 | 12 | BG041679 | BG041679 sv37e08.y |
| 37 | 936 | 42.7 | 661 | 14 | BQ863196 | BQ863196 GCG23D15.y |
| 38 | 934 | 42.7 | 814 | 10 | BE194793 | BE194793 HVSMEH008 |
| 39 | 927 | 42.4 | 567 | 10 | AM584093 | AM584093 N210025e |
| 40 | 917 | 41.9 | 582 | 10 | AW931213 | AW931213 EST357056 |
| 41 | 911 | 41.7 | 562 | 13 | BM528573 | BM528573 sal67h06.y |
| 42 | 899 | 41.1 | 566 | 10 | BE405193 | BE405193 WHE1211_D |
| 43 | 893 | 40.8 | 556 | 10 | BE433601 | BE433601 EST400130 |
| 44 | 893 | 40.8 | 558 | 10 | BE433609 | BE433609 EST400138 |
| 45 | 888.5 | 40.6 | 733 | 12 | BE705725 | BE705725 sc02_02d0 |

ALIGNMENTS

RESULT 1
AY109096
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AY109096 2423 bp mRNA linear HTC 25-MAY-2002
Zea mays PC0119715 mRNA sequence.
AY109096
AY109096.1 GI:21212506
HTC.
Zea mays.
Zea mays.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2423)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2423)
AUTHORS Coe,E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
FEATURES
source
1. 2423
/organism="zea mays"
/db_xref="taxon:4577"
/db_xref="taxon:4577"
/clone="PC0119715"
/clone.lib="Maize Mapping Project/DuPont Consensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 630 a 521 c 619 g 648 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 1.96e-202 Length: 2423
Score: 1750.00 Matches: 323
Percent Similarity: 88.40% Conservative: 35
Best Local Similarity: 79.75% Mismatches: 47
Query Match: 80.02% Indels: 0
DB: 11 Gaps: 0

US-09-720-451-6 (1-408) x AY109096 (1-2423)

QY 1 AlAaYValIleValhThrGlnLeuPheYrAspThrAspIlePheLeuLysPheValAsn 20
DB 618 GCTGACCTTATGTGTACACAACTTTCTATGAAACGACATCTTCTCAAGTTTGTGAAT 677
QY 21 AspCYArgGlnIleGlyIleThrCysProIleValIleProGlyIleMetProIleAsn 40
DB 678 GACTGCCGCCAATGTGTATTAATCTCCCATTTGTCGGCATATATGCAATAATAC 737
QY 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
DB 738 TACAAGGTTTCTCGCGACTGAGTGGCTGCGAATAAGATACCTTCTGAGATCACT 797
QY 61 AlAlAlaLeuCluProIleLysAspAsnGluAlaValLysAlaIleGlyIleHisLeu 80
DB 798 GCTGACCTAGATCTATCAAAAGCAATGAGAGGCTGTAGACAAATATGGAATCCACCTT 857
QY 81 GlyThrGlnMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThr 100
DB 858 GGAACGAGATGTGCAAGAAATTTCTGCTACTGCGATTAAGACTTTTGACACTTTACACA 917
QY 101 LeuAsnMetCluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluLysLeu 120
DB 918 CTAAACATGACAAAGTCTGCTATAGGAATTTGATGAATCTTGATTAATATGAGAGTCC 977
QY 121 LysValSerArgSerLeuProThrPheValGlyProAlaAsnValPheArgValLysGluAsp 140
DB 978 AAGGTTTCAAGGCCATTACTTGGAGGCCAGGACATAATGTTTCCGTGTTAAAGGAGAT 1037
QY 141 ValArgProIlePheThrPheAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTyrP 160
DB 1038 GTTCGACCTATATTTCTGGGCCAACACACCAAGAGCTATCTTAAAGGACATTAGGTTGG 1097
QY 161 AspGlnTyrProHisGlyArgTyrPheLysSerCysAsnProSerTyrGlyAlaLeuSer 180
DB 1098 GATCACTATATCCCATGAGAGGTGGGTGATTTCTCGAACCATCATATGAGACACTTACT 1157
QY 181 AspTyrGlnPheMetArgProAlaArgAspLysLysLeuValGluIleThrPheVal 200

DB 1158 GACCACAGTTTCACAAAGACCAAGAGCCGCTGTAGAAAGCTCCAAAGAGATGGCTGTT 1217
QY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
DB 1218 CCACGAAATCTGTGGAGACATTAAGTAGCGCTTCACAAACTCTGTGTCAAGGAACTC 1277
QY 221 ArgSerAsnProTyrPheSerGlyLeuAspGlyLeuGlnProGlyTyrLysIleLeuAsnGlu 240
DB 1278 ACAAGCACCCATGTGTGCAATTTGGAGCGTCTTACACAGAGACAAAGATTTATGATGAC 1337
QY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsn 260
DB 1338 CAGTTGGTGAATATTAAACCAAGAGGTTTCTTACAAATTAAACACCAACCTGCTGAAT 1397
QY 261 GlyLysLysSerAspSerProThrValGlyTyrPheGlyArgProGlyTyrValTyrGln 280
DB 1398 GGAGGAATCCGACTCCCTACTGTGTGGGTGGGTGCTCGGAGGCTACAGGTTATCAG 1457
QY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
DB 1458 AAGGCTACCTCGAATTTCTTGGCCCAAGAGAAAGTTGGACCAACTTAATTGAGAAATC 1517
QY 301 LysAspArgThrSerLeuThrTyrMetAlaValAsnLysAspGlySerTyrPheLysSerAsn 320
DB 1518 AAAGCATTCCTCTCTCACTTACATTCCTGTGAACAAAGATGAGAAACATTTCTCAAT 1577
QY 321 ValGlyGlnThrAspValAsnAlaValThrTyrPheGlyValPheProAlaLysGluIleLe 340
DB 1578 ATTTCACCGAACCGCGTAAATGCTGTACGTGGGTGTGTTTCCCTGCCCAAGAGATTAATC 1637
QY 341 GlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAlaPheGluIle 360
DB 1638 CAGCCTACGTTGTGAGATCATGACAGTTTATGTGTTGGAAGAGACAGAACTTTGAATATC 1697
QY 361 TrpSerArgGlyTyrPheLysSerLeuTyrProLysAspGluAlaSerArgLysLeuValGlu 380
DB 1698 TGGACTCGGGGTGGGTGGATGATCTCCGTGAGGCTATTCATGACGAGGCTATCGTAG 1757
QY 381 GluValGlyLysSerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
DB 1758 AAGTTTCGAAGACCTTACTTGTGTCAGCTCTTACAAAGCACTACGTCAGGGGAC 1817
QY 401 LeuPheAlaValPhe 405
DB 1818 CTGTTGCTGCTTC 1832

RESULT 2
LOCUS B0165347 851 bp mRNA linear EST 25-APR-2002
DEFINITION EST611216 KYKC Medicago truncatula cDNA clone pVKC-8D7, mRNA sequence.
ACCESSION B0165347
VERSION B0165347.1 GI:20307667
KEYWORDS EST.
SOURCE
ORGANISM Medicago truncatula
barrel medic.
REFERENCE 1 (bases 1 to 851)
AUTHORS VanderBosch,K., Endre,G., Silverstein,K., Town,C.D., Van Aken,S., Uiterback,T., Cheung,F. and Fraser,C.M.
TITLE The Medicago truncatula 'klionone' set: ESTs selected and re-arrayed from various libraries
JOURNAL Unpublished (2002)
COMMENT Contact: VanderBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvanderb@cbs.umn.edu


```

|||||  |||  |||||||||||||||||||||||||||||||||||||||
Db CCTGGAGGCCACCACTAATGTTTCCGTGCAAGAGGATGTCGCCCTATATTCG 120
QY 147 AIAAaATrProlYsSerYrTlLeSerYrTlLeGlyTPrAsGlyTPrHsGly 166
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 121 GCCAAGAGCAAGAGGTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 167 AqTPrGlyAspSerCysAsnProSerYrGlyAlaLeuSerAspYrGlnPheMetAr 186
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 181 CGGTGGGGGATTCAGAGAACCATCATAGAGAGGAGGAGGAGGAGGAGGAGG 240
QY 187 ProArGAlaArGAspLysLysLeuValGluGluTTPAlaValProLeuLysSerValGlu 206
    |||  |||  |||||||  |||||||  |||||||  |||
Db 241 CCACGCGGGCGGTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 207 AspLleTyrGluArGpPheArGpLeuTyrCysLeuGlyLysLeuArGSerAnPProTPrSer 226
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 301 GACATTAATGAGCGGTTCTGTAACCTTCTGTAAGGAGGAGGAGGAGGAGGAGG 360
QY 227 GluLeuAspGlyLeuGlnProGluTPrLysLleLleAsnGluGlnLeuGluLysLleAsn 246
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 361 GAGTTAGATGAGCTTCAACCCGAGACAGATATGATGATGATGATGATGATGATG 420
QY 247 ThLysGlyPheLeuThrLleAsnSerGlnProAlaValAsnGlyLysLysSerAspSer 266
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 421 TCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 267 ProThrValGlyTPrGlyLysGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhe 286
    |||  |||  |||||||  |||||||  |||||||  |||
Db 481 CTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 287 PheCysSerLysGlyLysLeuAspAlaLeuValAspLysCysLysAspYrTPrSerLys 306
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 541 TTCTGGCTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 307 ThTyrMetAlaValAsnLysAspGlySerTPrLysSerAsnValGlyLysThrAspVal 326
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 601 ACATACATCCGCGTGAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 327 AsnAlaValThrTPrGlyValPhe-ProAlaLysGluLleLleGln-ProThrLleValA 346
    |||  |||  |||||||  |||||||  |||||||  |||
Db 661 AACGCTGCTACTGGGGCGGTCCCGGGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 346 sPrProValSerPheAsnValTPr-LysAspGlyAlaPheGluLleTPrSerArGlyTPr 365
    |||  |||  |||||||  |||||||  |||||||  |||
Db 721 ACTGGGAGGAGCTTATGCTTGGAAAAAGATGAACGTTTACATCTGCTCAGGAGTGC 780
QY 366 Ala-SerLeuTyrProGluAspGluAlaSerArGlyLysLeuValGluGluValGlyLys 385
    |||  |||  |||||||  |||||||  |||||||  |||
Db 781 GCCCGCTGCTCCCGAAGGAGGATCATCCAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 385 rHisPheLeuValSerLeuValAspAsnAspYrTlLeAsnGlyAspLeuPhe 402
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 841 CTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891

```

RESULT 5
 BG418616
 LOCUS
 DEFINITION
 HVSMEK0023106f Hordeum vulgare testa/pericarp EST library
 HVCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0023106f, mRNA
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare.
 Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 ; Triticeae; Hordeum.
 1 (bases 1 to 823)
 REFERENCE
 AUTHORS
 Ming, R., Close, T. J., Kleinbols, A., Wise, R., Kannangara, G., von
 Weistein, D., Akhunov, E., Chin, A., Choi, D. W., Fenton, R. D., Kianian
 , P., Otto, C., Simons, K., Zhang, D., Begum, D., Frisch, D., Yu, Y.,

TITLE
 JOURNAL
 COMMENT
 Henry, D., Palmer, M., Rambo, T., Simons, J., Oates, R. and Main, D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Morex testa/pericarp cDNA library
 Unpublished (2001)
 Contact: Ming RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 590
 Seq primer: AATTAAACCTCAGCAAGG
 High quality sequence stop: 759.
 Location/Qualifiers
 1..823
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMEK0023106f"
 /clone_lib="Hordeum vulgare testa/pericarp EST library
 HVCDNA0013 (normal)"
 /tissue_type="testa/pericarp"
 /lab_host="TTC121"
 /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
 plants were raised from seeds in a controlled environments
 growth chamber maintained in continuous light at 18°C, and
 testa and pericarp were dissected from developing kernels
 at Washington State University, Pullman, WA (Kannangara,
 von Weistein). Total RNA was prepared, poly(A) RNA was
 purified, one cDNA library was made, and 1 million plu
 were in vivo excised to give plasmid SK(-) cDNA
 phagemids in the T7 Close lab at the University of
 California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
 Kianian, Otto, Simons, Zhang), phagemids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Ming). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Ming, Yu, Frisch, Henry, Simons, Oates
 , Rambo, Main). The sequence has been trimmed to remove
 vector sequence and contains a minimum of 100 bases of
 phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Ming R, Kleinbols A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/g9pages/bgr/31/cover.html)"

BASE COUNT 236 a 187 c 207 g 192 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4.59e-127 Length: 823
 Score: 1128.00 Matches: 213
 Percent Similarity: 85.31% Conservative: 23
 Best Local Similarity: 77.17% Mismatches: 38
 Query Match: 51.58% Indels: 2
 DB: 12 Gaps: 1

US-09-720-451-6 (1-408) x BG418616 (1-823)

```

QY 68 ASPASnGluGluAlaValLysAlaTyrGlyLleHisLeuGlyThGluMetCysLysLys 87
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 2 GACAAATGAGAGAGGCTGTGAAGAGCATATGGAATCCACTTGTACTGAGATGCAAGAA 61
QY 88 lLeuAlaHisGlyLleLysThrLeuHisLeuTyrThrLeuAsnMetGluLysSerAla 107
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 62 AATTTGGCTAGTGGGATCAAGACTTTCACCTTACACTTAACATGAAAAAGCCGT 121
QY 108 LeuAlaLleLeuMetAsnLeuGlyLleGluGluSerLysValSerArGserLeuPro 127
    |||||||  |||||||  |||||||  |||||||  |||||||  |||
Db 122 TTAGCAATTCGTGATGATCTGGAATTAATAGAGGAGGAGGAGGAGGAGGAGGAGG 181

```


|||||
Db 184 CTTGACGACAGCAAGATATATAGTCTGCTGAAAAGATTACACAAAGCGCTTC 243
Qy 251 LeuThrIleAsnSerGlnProAlaValAsnGlyIuLysSerAspSerProThrValGly 270
Db 244 CTCACCAATCAATACGACGCGCTCAATGGGAAAAGTCAGATCTCTCTACTGTTGGC 303
Qy 271 TTPGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 304 TGGGTGGACACAGGGGGTATGTTTACCAGAAGCATGTAGAGATTCTCTGCTCAAG 363
Qy 291 GluIysLeuAspAlaLeuValAspLysCysLysAspArgTyrSerLeuThrTyrMetAla 310
Db 364 GAAAAGCTGATGCACCTTGATATAAATGCAAGATCGAACATCTTTAACTTAATAGCT 423
Qy 311 ValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAsnAlaValThr 330
Db 424 GTGAATTAAGACGGAGAGCTGGAAATCTAATGTGGGTCAAACTGATGTAATGCTGTGACA 483
Qy 331 TTPGlyValAlaPheProAlaLysGluIleIleGlnProThrIleValAspProValSerPhe 350
Db 484 TGGNGTGTCTCCAGCTAAGAGATATATCAACCAACCATGTGTGATCTGTGAGCTTC 543
Qy 351 AsnValTTPLysAspGluAlaPheGluIleTyrSerArgGlyTTPAlaSerLeuThrPro 370
Db 544 AATGATGGAAGATGANCATTTGATAAATTTGGTCAAGAGGAGGCGCACTTGCCCT 603
Qy 371 GluAspGluAlaSerArgLysLeuValGluGluValGlyGlySerHisPheLeuValSer 390
Db 604 GAGGGGATGCATTCACGAAATTTGGTTGAAGAATTTGGGGGACGCTTACTTCTGTGAGT 663
Qy 391 Leu 391
Db 664 TTG 666

RESULT 7
BO994876 760 bp mRNA linear EST 21-AUG-2002
LOCUS QGF8G21.yg.ab1 QG.EFGHJ lettuce serriola Lactuca sativa cDNA clone
DEFINITION QGF8G21.yg.ab1 QG.EFGHJ lettuce serriola Lactuca sativa cDNA clone
ACCESSION BO994876
VERSION BO994876.1 GI:22414411
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
1 (bases 1 to 760)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J., Ellison
,P., Kojkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://comgenomics.ucdavis.edu/
unpublished (2002)
Contact: Alexander Kozik (R.W.Michelmore)
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@ucdavis.edu]
belongs to contig QG_CA.Contig1724, see http://cgpdb.ucdavis.edu/
for details.
Plate: QGF8 row: G column: 21.

FEATURES
Source location/Qualifiers
1..760
/organism="Lactuca sativa"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGf8G21"

/clone.lib="QG.EFGHJ lettuce serriola"
/lab.host="E.coli"
/note="Vector: pBRCDNASFIAB. The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB-QG.EFGHJ lettuce serriola
TAG_TISSUE-flowers post-fertilized
TAG_SEQ="GCCATCGGC"

BASE COUNT 226 a 167 c 175 g 192 t
ORIGIN

Alignment Scores:
Pred. No.: 2,516-123 Length: 760
Score: 1097.00 Matches: 198
Percent Similarity: 88.89% Conservative: 26
Best Local Similarity: 78.57% Mismatches: 28
Query Match: 50.16% Indels: 0
DB: 14 Gaps: 0

US-09-720-451-6 (1-408) x BO994876 (1-760)

Qy 35 IleMetProIleAsnAsnTyrLysGlyPheIleAlaArgMetThrGlyPheCysLysThrLys 54
Db 3 ATCATGGCCCAATTAACAACCTTAATGATGTTCAATGAGATGCTGTTCTTCCAAAATAAG 62
Qy 55 IleProAlaAspIleMetAlaAlaLeuGluProIleLysAspAsnGluAlaValLys 74
Db 63 ATTCATCTGTAATTTACTGCTGCTTGGAGCCCTATCAAGACAAATGAGAGCTGTTA 122
Qy 75 AlaTyrGlyIleHisLeuGlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLys 94
Db 123 AACTATGAGATGCCACCTTGGAACTGAAATGTGTAGAATAATCAAGGAGCTGATCAAG 182
Qy 95 ThrLeuHisLeuTyrThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeu 114
Db 183 ACTTTGATCTTTTATACCTCACTACACATGACAAATCTGCATTTGGCAATTTGACAAAGCT 242
Qy 115 GlyLeuIleGluGluSerLysValSerArgSerLeuProTTPArgTTPProAlaAsnVal 134
Db 243 GGATTTGTTGACGAGGCCAAATTAACAAGCACCTTCATGAGAGACCTACTAACCT 302
Qy 135 PheArgValLysGluAspValArgProIlePheThrPAlaAsnArgProLysSerTyrIle 154
Db 303 TTCCGCTTTAAAGAAGATGTAGGCCAATTTTGGGCCAATCGTCCAAAGAGCTACATA 362
Qy 155 SerArgThrIleGlyTTPAspGlnTyrProHisGlyArgTTPGlyAspSerCysAsnPro 174
Db 363 TCAAGGACGGGTGTGGAGCAAAATACACATGGCGCTTGGGGGATTTCTTAAGCA 422
Qy 175 SerTyrGlyAlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLeu 194
Db 423 TCATAGGAGCATTTAACCGATCATCATGAGCCACGTCAGTCAAGAAATTT 482
Qy 195 ValGluGlnTTPAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeu 214
Db 483 CAAGCAAGATGGGTGCTTGTGACCAAGTTTGATGATATTTCCGAAATTTCTGAAG 542
Qy 215 TyrCysLeuGlyLysLeuArgSerAsnProTTPSerGluLeuAspGlyLeuGlnProGlu 234
Db 543 TACTGCTTGGGCACTATAAAACAGCCCTTGTGGACCTGAGAGCCTTCAAGCAAG 602
Qy 235 ThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsn 254
Db 603 ACAAAATCATCAACGAGAGACTGGCTTCCATCAACCTCAAGGTTTCTCAACATCAAC 662
Qy 255 SerGlnProAlaValAsnGlyLysSerAspSerProThrValGlyTTPGlyLysPro 274

ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 700)
AUTHORS Cushman, J. C.
TITLE An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
JOURNAL Unpublished (1997)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR primers
FORWARD: T7
BACKWARD: T3
Plate: LD-16 row: H column: 4
Seq primer: T3
High quality sequence stop: 350
POLYA-No.

FEATURES
source location/Qualifiers
1..700
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L0-1588"
/clone_lib="Ice plant Lambda Uni-Zap XR expression library
, 0 hours NaCl treatment"
/tissue_type="leaf"
/dev_stage="Six week old"
/note="Vector: lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 226 a 136 c 165 g 173 t
ORIGIN

Alignment Scores:
Pred. No.: 3.66e-122 Length: 700
Score: 1087.00 Matches: 198
Percent Similarity: 92.70% Conservative: 18
Best Local Similarity: 84.98% Mismatches: 17
Query Match: 49.70% Indels: 0
DB: 10 Gaps: 0

US-09-720-451-6 (1-408) x BE576981 (1-700)

QY 48 ThrGlyPheCysLysThrLysIleProAlaSerIleMetAlaIleLeuGluProIleLys 67
|||||
Db 1 ACTGTTCTCTGTAACAACAAGATTCTCAAGAGTTACTGATGCATTGGAACCAATTAG 60
|||||

QY 68 AspAsnGluGluAlaValLysAlaTyrGlyIleHisLeuGlyThrGluMetCysLysLys 87
|||||
Db 61 GACAAATGATGAGGCCGTGAGGCCCTACGGGATTCCACCTCGGAACATGGAATGTGCAAGAG 120
|||||

QY 88 IleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrIleAsnMetGluLysSerAla 107
|||||
Db 121 ATTTGGCTAGTGAGATTAAACAGTTGCAATGATATACCTCGAATATGACAAACTCA 180
|||||

QY 108 LeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSerLysValSerArgSerLeuPro 127
|||||
Db 181 TTAGCAATTTTGACGAATCTTGATTAATTGAAATCCAAATCCAAAGGTCAATTACT 240
|||||

QY 128 TTPATGATGProAlaAsnValPheArgValLysGluAspValArgProIlePheTrpAla 147
|||||
Db 241 TGGAAACGCCCTGCAAAATGTGAACCGTTAAAGAGATGTTGTCATATTTTGGGCA 300
|||||

QY 148 AsnArgProLysSerTyrIleSerArgThrIleGlyTTPASPGLITyrProHisGlyArg 167
|||||
Db 301 AATTCGTCAAAAGTTACTTACGGGACAGTTGATGGATGATACCAACAGGGGGA 360
|||||

QY 168 TTPGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyrGlnPheMetArgPro 187
|||||

Db 361 TGGGATGATTCGCAATGCTTCTTATGAGGACATTAGTACCATCAGTTATGAGGCCA 420
QY 188 ArgAlaArgAspLysLysLeuValGluGluTTPAlaValProLeuLysSerValGluAsp 207
|||||
Db 421 CGTCCCGGATGAAGAAGCTCCAAAGAAGAGTGCTGCCACTAAAAGTGTGAAGAT 480
QY 208 IleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTyrSerGlu 227
|||||
Db 481 ATACATGAGAGGTTCCAAACTACCTCTGTTAGCTTAAGCTAAAGTATCCTGCTGAA 540
QY 228 LeuAspGlyLeuGluProGluThrLysIleLeuAsnGluLeuGluLysIleAsnThr 247
|||||
Db 541 ATGATAGCGCTTACGACAGAAACAGATCATCAATGAACAGCTG6CCAAATATAAG 600
QY 248 LysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspPro 267
|||||
Db 601 AAGGCTTCTCACCATCAACACCACTGCTGTTAAATGAGAAAATCTAGTCCCA 660
QY 268 ThrValGlyTyrGlyLysProGlyGlyTyrValIleTyrGln 280
Db 661 AGTGTGATGGGAGAGCTCTGATGATATATACCA 699

RESULT 10
BI421654
LOCUS EST532320 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION CLEC67D19 5' end, mRNA sequence.
ACCESSION BI421654
VERSION BI421654.1 GI:15195373
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 752)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E.,
'Liang, F., Upton, J., Craven, M. B., Bowman, C. L., Ahn, S., Roming,
'C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and Giovannoni, J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
location/Qualifiers
1..752
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEC67D19"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
undifferentiated masses. Tomato Callus EST Library"

FEATURES
source
1..752
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEC67D19"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 216 a 164 c 185 g 187 t
ORIGIN

Alignment Scores:
Pred. No.: 3.42e-121 Length: 752
Score: 1079.50 Matches: 204
Percent Similarity: 89.24% Conservative: 20
Best Local Similarity: 81.27% Mismatches: 27
Query Match: 49.36% Indels: 1
DB: 13 Gaps: 0

US-09-720-451-6 (1-408) x B4121654 (1-752)

QY 88 ILELEUALHISGLYIILEYSTRHLEUHSLEUTYRTHRLAUSMETSGLYLSERALA 107
 DB 2 ATTTTAGCCATGCGCTGACAGACCTTGTATACCTTAACATGAGAAATTCACCA 61

QY 108 LEUALAILELEUMETASNLLEUGLYLEULIEGLUGLUSERLYVALSERARSERLEUPRO 127
 DB 62 TTGGCCATTTTGTGATCTTGATTAATCGAAGATCCAAATTTCTAGGCCATTGGCT 121

QY 128 TTPATGATGPROALAASNLVALPHEARGVALLYSGUASPVALARGPROILEPHEPTRPALA 147
 DB 122 TGGAGACGTCTGCACAAATVTTTTCCGTCAAGAGAGATGCTCCTATATTGTTGGGCG 181

QY 148 AANATGPROLYSSEYTRILESERARGTHRIIEGLYTPASPGLNTPROHISGLYARG 167
 DB 182 AATCGCCAAAGAGCTACATTTCAGAGACACAGGTGGAGATGATACCCACATGTTAGA 241

QY 168 TTPGLYASPSEYCYSHNPROSEYTRGLYALALEUSERASPTPYGLNPHEMETARGPRO 187
 DB 242 TGGAGCAATGCTCAAAATTCATCATACGAGCATTTACCGACTATCAGTTCATGCGTCA 301

QY 188 ARGALATARGASPLYSLEUVALGLUGLUTTPALAVALPROLEULYSERYALGLUNSP 207
 DB 302 CGCTCAAGAGATGAGAGCTTCAAGAGAGATGCGCTGCGCTTGAAATGCGTGAGAGAT 361

QY 208 ILETYGLUARGPHEARGLEUTRYCYLSLEUGLYLYLSLEARGSERANPROTRPSERGLU 227
 DB 362 ATCTGACGAGAAATTCAGAGACTGCTGTGAAAGCTGAGAGAGTGGCCATGCTTGA 421

QY 228 LEUASPGLYLEUGLNPROGLUTHRILYSILELEASNGULGNLEUGLYLSILEASNTHR 247
 DB 422 TTAGATGGCTTCCACCGACAGACAAAGATCAATAACGAAAGTTGGTCAACACAG 481

QY 248 LYSGLYPHELEUTHRIEASNSERGLNPROALAVALSNGLYLSLEUSERSASPSERPRO 267
 DB 482 AAGGCTTCTCTACTATTAACAGCCACAGACAGATTAATGCTGAGAAATCGACCTCTCT 541

QY 268 THRVALIYTPGILYLPYPROGLIYLYTYRVALTYRGLNLYSALATYRVALIUPHEPHE 287
 DB 542 TCTGTTGGATGGGTGGCCCGCGGATGTTTATCAAAAGGCTACTTGAGATCTCTTC 601

QY 288 CYSSEYRGLYLSLEUVALSPALALEUVALSPLYSCYLSYASPAVGYTHSERLEUTHR 307
 DB 602 TGTCTTCAGGAGAAATTTACGCCCTCTGTGAGAAATGCAAGTCTTTCCTTACCCCTCA 661

QY 308 TYRMETALAVALSNLSPYSPYSEYTRPYSSERASVALGLYGLNTHRASPVALASN 327
 DB 662 TACATGGCGCTGTACAGAGAGAGGAACTGGATCTTCAACGTCA--CCAAACGACATCAAC 719

QY 328 ALAVALTHRTTPGILYVALPHEPROALALYSGLU 338
 DB 720 GCTGTGACATGGGAGTTTCCACAGCTAAGAA 752

RESULT 11
 LOCUS B418214 842 bp mRNA linear EST 23-OCT-2001
 DEFINITION HYDMEK0021017f Hordeum vulgare testa/pericarp EST library
 HVCNMA0013 (normal) Hordeum vulgare cDNA clone HYDMEK0021017f, mRNA
 sequence.
 ACCESSION B418214
 VERSION B418214.1 GI:13323765
 KEYWORDS EST.
 SOURCE Hordeum vulgare.
 ORGANISM Hordeum vulgare.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 842)
 REFERENCE Wang R., Close, T.T., Klein, H.A., Wise, R., Kannangara, G., von
 Weizstein, D., Akhunov, E., Choi, D.W., Fenton, R.D., Kianian,
 P., Otto, C., Simons, R., Zhang, D., Begum, D., Frisch, D., Yu, Y.,

TITLE Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and Main, D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Morex testa/pericarp cDNA library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Ming RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu
 Total hg bases = 539
 Seq primer: AATTAACCTCACTAAGG
 High quality sequence stop: 641.
 location/Qualifiers
 1..842
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HYDMEK0021017f"
 /clone_1lb="Hordeum vulgare testa/pericarp EST library
 HVCNMA0013 (normal)"
 /tissue="testa/pericarp"
 /lab_host="TUC121"
 /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 Plants were raised from seeds in a controlled environments
 growth chamber maintained in continuous light at 18°C, and
 testa and pericarp were dissected from developing kernels
 at Washington State University, Pullman, WA (Kannangara,
 von Weizstein). Total RNA was prepared, poly(A) RNA was
 purified, one cDNA library was made, and 1 million pfu
 were in vivo excised to give plasmid SK(-) cDNA
 phagmids in the rt close lab at the University of
 California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
 Kianian, Otto, Simons, Zhang). Phagmids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Wang). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wang, Yu, Frisch, Henry, Simmons, Oates
 , Rambo, Main). The sequence has been trimmed to remove
 vector sequence and contains a minimum of 100 bases of
 phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see close rt, Wang R., Klein, H.A., Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/dgn/31/cover.html)"
 BASE COUNT 238 a 193 c 215 g 195 t 1 others
 ORIGIN

Alignment Scores:
 pred. No.: 8,39e-118 Length: 842
 Score: 1052.50 Matches: 209
 Percent Similarity: 84.25% Conservative: 21
 Best Local Similarity: 76.56% Mismatches: 42
 Query Match: 48,13% Indels: 3
 DB: 12 Gaps: 1

US-09-720-451-6 (1-408) x B418214 (1-842)

QY 70 GLUGLALAVALSALATYRGLIENISLEUGLYTHRGUMETCYLSYSTLEULEU 89
 DB 2 GAGGAGGCTGTCAAAACATATGGAATCCACTTGTGACATGAGATGGCAAGAAATTTTG 61

QY 90 ALAHISGLYIILEYSTRHLEUHSLEUTYRTHRLAUSMETSGLYLSERALA 109
 DB 62 GCTAGTGGGATCAAGACTTGGACCTTTACACACTTAACATGAGAAAGCCCTTTAGCA 121

QY 110 ILELEUMETASNLLEUGLYLEULIEGLUGLUSERLYVALSERARSERLEUPTRPARG 129
 DB 122 ATTTGATGAACTTGATTAATGAGGAATCCAAAGCTTTCAAGAAACATTTACCTTGAGG 181

130 ArgProAlaAsnValPheArgValIleuAspValArgProIlePheTrpAlaAsnArg 149
 182 CCACCACTAATGTTTCCGTCGAAGAGAGATGTCGCCCTATATCTGGCCACACAG 241
 150 ProLysSerTrpIleSerArgThrIleGlyTrpAspGlnTrpProHisGlyArgTPGly 169
 242 CCAAGAGTACATTTCAAGAGCACCTGGTGGATCAATACCCACAGAGGGCGGTGGGT 301
 170 AspSerCysAsnProSerTrpGlyAlaIleuSerAspTrpGlnPheMetArgProArgAla 189
 302 GATTCACAGAACCCATCATACGAGCCCTTATATACACACAGTTCAACAGCCACGTCGG 361
 190 ArgAspLysIleuValGluGlnTrpAlaValProLeuLysSerValGluAspLysTrp 209
 362 CCGTGAAGAAGCTCCAGAGGAATGGCTGTCCACAGAAATCTGTGCAAGACATTAAT 421
 210 GluArgPheArgLeuTrpCysLeuGlyLysLeuArgSerAsnProTrpSerGluLeuAsp 229
 422 GAGCGGTTCTGTAACCTCTGTGAAGAAACTTAAAGCACCCATGATGCTGAGTTAGAT 481
 230 GlyLeuGlnProGluTrpLysIleIleAsnGlnIleuGlnLysIleAsnTrpLysGly 249
 482 GGTCTTCAACCCGAGACAAAGATATGATGATGATGATGATGATGATGATGATGATGAT 541
 250 PheLeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrVal 269
 542 TTCTTACCATCATACAGCCACCTCTGTAATGCAAGAAATCTGATCTCTCATNTGTT 601
 270 GlyTrpGlyGlyProGlyLysTrpValTrpGlnLysAlaTrpValGluPheCysSer 289
 602 GATGGGGGGGCTCCAGAGAGCTATGTTACCA -AAGGCTACGTCGATTTCTTGGCGCT 660
 290 LysGluLysLeuAspAlaLeuValAspLysCysLysAspArgTrpSerLeuThrTyrMet 309
 661 AAGGACAGAGTGGACCACTCATCGAAGAAAGAGAGGCAATCTTCCCTCATCATCATC 720
 310 AlaValAsnLysAspLysSerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
 721 GCGGGGACAGAGGAGGAATGATCTCAACATCCCTGCAAGAGCCG---GGAGCTGT 777
 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnPro 342
 778 CTTGGGGCGCTGTTCCCGG -CAAGAAATCATCCACCT 815

RESULT 12
 BE805904 611 bp mRNA linear EST 06-DEC-2001
 LOCUS ss62h07.y1 Gm-c1062 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-c1062-1046 5' similar to SW:MTNR_ARATH 080585 PROBABLE
 METHYLENETETRAHYDROFOLATE REDUCTASE ; mRNA sequence.
 BE805904
 ACCESSION BE805904.1 GI:10237016
 VERSION
 KEYWORDS soybean.
 SOURCE Glycine max
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 611)
 REFERENCE
 AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

FEATURES
 source
 1. 611
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1062-1046"
 /tissue_type="stem tissue of greenhouse grown plants"
 /dev_stage="1 month old"
 /lab_host="DH10B"
 /note="Vector: Bluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from stem tissue of 1 month old greenhouse grown plants
 for the cultivar Raiden. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with a XhoI restriction site. EcoRI
 adapters were ligated to the blunt-ended cDNA fragments
 followed by XhoI digestion. The cDNA fragments were
 directionally cloned into the EcoRI-XhoI restriction site
 of the Bluescript vector. The ligated cDNA fragments were
 transformed into DH10B host cells (GibcoBRL). This library
 was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 178 a 101 c 165 g 166 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 5,61e-118 Length: 611
 Score: 1052.00 Matches: 196
 Percent Similarity: 98.03% Conservative: 3
 Best Local Similarity: 96.35% Mismatches: 4
 Query Match: 48.10% Gaps: 0
 DB: 12

US-09-720-451-6 (1-408) x BE805904 (1-611)

QY 202 LeuLysSerValGluAspIleTrpGluArgPheArgLeuTrpCysLeuGlyLeuArg 221
 DB 3 TTGAAAAGCATTCAGATATATATAGAGAGTTTAACTGATGTCTTGAAGATTGAGA 62

QY 222 SerAsnProTrpSerGluLeuAspGlyLeuGlnProGluTrpLysIleAsnGluGln 241
 DB 63 AGCAATCCTTGTCACAGACTAGATGCTTCAGCCAGACAAAGATATATAGAGCTG 122

QY 242 LeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAsnGly 261
 DB 123 CTGGAAAAGATTAAACAAAGGGCTTCACCATCAATAGCCACAGCTGTCAATGG 182

QY 262 GluLysSerAspSerProThrValGlyTrpGlyGlyProGlyLysTrpValTrpGlnLys 281
 DB 183 GAAAAGTCAATTTCTCTTACTGTGGCTGGGAGGACAGGGGAGTTTACAGAGA 242

QY 282 AlaTrpValGluPhePheCysSerLysGlyLysLeuAspAlaLeuValAspLysCysLys 301
 DB 243 GCATATGATAGAGTCTCTCTCTTAAGGAAAGCTGAGGACCTGTTATTAATGCAAG 302

QY 302 AspArgThrSerLeuThrTrpMetAlaValAsnLysAspGlySerTrpLysSerAsnVal 321
 DB 303 GATCAACATCTTAACTTATATAGCTGTGATTAAGAGGAGGAGCTGGAATCTATGTG 362

QY 322 GlyGlnThrAspValAsnAlaValThrTrpGlyValPheProAlaLysGluIleGln 341
 DB 363 GGTCAAACTGATGTTATGCTGAGACATGGGCTGTCCAGCTTAAGAGATTAATTCAA 422

QY 342 ProThrIleValAspProValSerPheAsnValTrpLysAspGlnAlaPheGluIleTrp 361
 DB 423 CCAACCATTTGATCTCTGTCAGCTTCATATGATGAAGAGATGAGCATTTGAAATTTGG 482

QY 362 SerArgGlyTrpAlaSerLeuTyrProGluAspGluAlaSerArgGlyLeuValGluGln 381
|||||
Db 483 TCAGAGAGATGGGCAAGCTTGTACCTGATGCGTATGATCCAGCAAAATTTGGTTGAAGAG 542
QY 382 ValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrTrpIleAsnGlyAspLeu 401
|||||
Db 543 GTGGGGGACGACTTCTTGGTGTGATGTCGACATGATATCATCAATGATGATCTT 602
QY 402 PheAlaVal 404
|||||
Db 603 TTCACCGCTC 611
RESULT 13
BG123885
LOCUS BG123885 729 bp mRNA linear EST 31-JAN-2001
DEFINITION EST469531 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOF3B13 5' sequence, mRNA sequence.
ACCESSION BG123885.1 GI:12624073
VERSION BG123885.1 GI:12624073
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 729)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
Hansen,C., Rohnig,S. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
FEATURES
source
1..729
/organism="Lycopersicon esculentum"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="CTOF3B13"
/clone_11b="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI. Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA96). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 216 a 155 c 180 g 178 t
ORIGIN
Alignment Scores:
Pred. No.: 1,76e-117 Length: 729
Score: 1049.00 Matches: 200
Percent Similarity: 89.75% Conservative: 19
Best Local Similarity: 81.97% Mismatches: 22
Query Match: 47.97% Indels: 3
DB: 12 Gaps: 0
US-09-720-451-6 (1-408) x BG123885 (1-729)
QY 101 LeuAsnMetGluIysSerLeuAlaLeuMetAsnLeuGlyLeuIleGluGlnSer 120
Db 3 TTCAACATGAGCAATACACATTTGATGAAATCTTGATTAATCGAAGAGCTCC 62
QY 121 LysValSerArgSerLeuProTrpArgArgProAlaAsnValPheArgValLysGluAsp 140
|||||
Db 63 AAAATTTTCAAGCCATTGCTTGAGAGAGCTCTCGCAAAATATTTCCTGTCAAGAGAGAT 122
QY 141 ValArgProIlePheTrpAlaAsnArgProLysSerTyrTrpIleSerArgTrpIleGlyTrp 160

Db 123 GTGGCTCTATATTTTGGGCGAATGCTCCAAAGACCTACATTTCAAGACCAAGCTTG 182
QY 161 AspGlnTyrProHisGlyArgTrpGlyAspSerGlyAsnProSerTyrGlyAlaLeuSer 180
Db 183 GATCAATATCCACATGCTATGATGAGCAAAATGCTCAAAATCATCATACGAGCACTTACC 242
QY 181 AspTrpGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGluTrpAlaVal 200
Db 243 GACTATCATGTTCAATGCGTCAGCGCTCAAGAGATTAAGAAGCTTCAAGAGATGGCTGCG 302
QY 201 ProLeuLysSerValGluAspLysIleTyrGluArgPheArgLeuTyrCysLeuGlyLysLeu 220
Db 303 GCTTTGAATAGCGTGAAGATATCTACGAAATTCAGAACTCTGCTTGGAAAGCTG 362
QY 221 ArgSerAsnProTrpSerArgGluLeuAspGlyLeuGlnProGluTrpLysIleIleAsnGln 240
Db 363 AGAAGCTGCCCATGCTGCTAGTATGATGCTTACGCCAGAGCAAGATCATTAACGAA 422
QY 241 GlnLeuGluLysIleAsnTrpLysGlyPheLeuTrpIleAsnSerGlnProAlaValAsn 260
Db 423 AAGTTGGGTGAGGTCAACAGCAAGAGCTTCTGACTATTACAGCAACGAGCTTAAT 482
QY 261 GlyLysLysSerAspSerProTrpValGlyTrpGlyGlyProGlyGlyTyrValTrpGln 280
Db 483 GCTGAGAAGTCCACACTCTCTCTGTTGATGGGTGGCCCGCGGATATGTTATCA 542
QY 281 LysAlaTrpValGlnPhePheCysSerLysGluLysLeuAspAlaLeuValAspLysCys 300
Db 543 AAGGCATATCCCTTACTACTTCTTCTGCTTCGAGAGATGAAAGCCCTGTTGAAATG 602
QY 300 sLysAspArgTrpSerLeuTrpTyrMetAlaValAsnLysAspGlySerTrpLysSer 320
Db 603 CAAGTCTTCCCTTACTACTTACTTACTGAGCTGTAAACAGAGAAAGAAACGATCTTCA 662
QY 320 snValGlyGlnTrpAspValAsnAlaValTrpTrpGlyValPheProAlaLysGluIle 339
Db 663 ACGTCAACCAACGACATCAACGCTGTGACATGGGGAGTTTCCCGCTTAAGGAATT 722
QY 340 IleGln 341
Db 723 ATACAA 728
RESULT 14
BE643238 821 bp mRNA linear EST 01-SEP-2000
LOCUS BE643238
DEFINITION Cri2_8_E11.SP6 Ceratopteris Spore Library Ceratopteris richardii
cDNA clone Cri2_8_E11 5', mRNA sequence.
ACCESSION BE643238
VERSION BE643238.1 GI:9960923
KEYWORDS EST.
SOURCE Ceratopteris richardii.
ORGANISM Ceratopteris richardii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Filicophyta; Filicopsida; Filicales; Pteridaceae; Ceratopteris.
REFERENCE 1 (bases 1 to 821)
AUTHORS Chatterjee,A., San Miguel,P., Stout,S.C., Banks,J., Ceratopteris,
TITLE Expressed sequence tags of cDNA clones from a C. richardii library
JOURNAL Unpublished (2000)
COMMENT Contact: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: stroux@utmsi.utexas.edu
Plate: Cri2_8 row: E column: 11
Seq primer: SP6.
FEATURES
source
1..821
/organism="Ceratopteris richardii"
/cultivar="Brogn"
/db_xref="taxon:49495"

```

/clone="Cri2_8_E11"
/clone.lib="Ceratopteris Spore Library"
/tissue.type="Gametophyte"
/cell.type="Spore"
/dev.stage="20 hours after germination initiation"
/notes="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."
BASE COUNT      256 a      162 c      187 g      214 t
ORIGIN
Alignment Scores:
Pred. No.:      6,66e-117      Length:      821
Score:          1045.00      Matches:      188
Percent Similarity: 83.52%      Conservative: 40
Best Local Similarity: 68.86%      Mismatches: 45
Query Match:    47.78%      Indels:      0
DB:              10      Gaps:      0

US-09-720-451-6 (1-408) x BE643238 (1-821)
QY      49 GlyphecyslysthrlyleProalaaspIleMetAlaLeuGluProIleLysasp 68
      |||||||
      1 GGCTTTGCAAGACTTAAGATTAAGTCTTACCAAAATGCGTGAGAGCTTCAAGAT 60
QY      69 AsnglUGLAlaValAlaTyrGlyLeuHleuGlyThrCluMetCysLysLysIle 88
      |||||||
      61 GATGACAGAGCTGTACACCATATGATATGATGATGATGATGATGATGATGATGAT 120
QY      89 LeuAlaHleuGlyLeuHleuHleuHleuHleuHleuHleuHleuHleuHleu 108
      |||||||
      121 TTGGCTCATGCGTAAAGATCGCATTTACACCTTACCTGGAATAATCAGCTGTG 180
QY      109 AlalleuMetLysLeuGlyLeuGlyLeuGlyLysValSerArgSerLeuProtrp 128
      |||||||
      181 GGTATCTGACAACTTGGCTTGTGATATTTCAAAATAGCGGCGCATTCGACG 240
QY      129 ArgArgProAlaAsnValPheArgValLysGluAspValArgProIlePheTrpAlaAsn 148
      |||||||
      241 AGGCTTCAGCAATATCAACAGAGATTAGAGAGATTTGGCCATTTTCTGGCCAT 300
QY      149 ArgProLysSerTyrIleSerArgThrIleGlyTrpAspGlnTyrProHleuGlyArgTrp 168
      |||||||
      301 CGTCTCGAGATTACATCTCTAGACAGAGACCTGGAGAGATTTCCCTCGGAGAGTGG 360
QY      169 GlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyrGlnPheMetArgProArg 188
      |||||||
      361 GGGACACTCGCAACCTGATATGAGGCTTATCTGATTCACAGTTCTTACGACGAGAG 420
QY      189 AlaArgAspLysLysLeuValGluGluTrpAlaValProLeuLysSerValGluAspIle 208
      |||||||
      421 TCTGCAACAAGAAATCCAAAGAGTGGTGTCTCTTGAATCAGATCAACGACATTT 480
QY      209 TyrGluArgPheArgLeuTyrCysLeuGlyLysLeuArgSerAsnProTrpSerGluLeu 228
      |||||||
      481 TATCAGTATTTGCCAATATTCCTTCAGGAGAAATTAAGACCAATCTTGGTCAAGCTT 540
QY      229 AspGlyLeuGlnProGluTrpLysIleLeuAsnGluGluLeuGluLysIleAsnThrLys 248
      |||||||
      541 GAAGCTTGGACGCAAGAACTGAAATCAATCAATTAATTTAGTAGAGATCAACCTAAA 600
QY      249 GlyPheLeuThrIleAsnSerGlnProAlaValAsnGlyLysLysSerAspSerProthr 268
      |||||||
      601 GGTCTTCTCAAAATTAATATGCTGCTGCTTAATGAGGAGAAATCAAGACAGCTCCACA 660
QY      269 ValGlyTrpGlyLysProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCys 288
      |||||||
      661 TATGCTTGGGAGAGCTGTGATGATATCTTACACAGAGCATATTAATAATTTTCTGT 720
QY      289 SerLysGluLysLeuAspAlaLeuValAspLysCysLysAspArgThrSerLeuThrTyr 308
      |||||||
      721 TCACGGGCAAAACTTGACAAACTCTGAGAGAAAGCAAAATCTTCCATCTTGAATAT 780

```

```

QY      309 MetAlaValAsnLysAspLysSerTrpLysSerAsnVal 321
      |||||||
Db      781 ATTGCCGTAAGCTCAAGAGGTGAGACTGACGCAATATTT 819

RESULT 15
BF071278
LOCUS
DEFINITION
  BF071278          588 bp      mRNA      linear      EST 06-DEC-2001
  st42e03.y1 Gm-c1067 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
  Gm-c1067-1734.5, similar to SW:MTNR_ARATH O80585 PROBABLE
  METHYLENETETRAHYDROFOLATE REDUCTASE ;, mRNA sequence.
ACCESSION
  BF071278
  BF071278.1      GI:10846202
VERSION
  EST.
KEYWORDS
  soybean.
SOURCE
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 588)
  Shoemaker,R., Kelm,P., Vodkin,L., Erpelidg,J., Corvett,V., Khana
  ,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
  ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
  ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
  ,R., Materston,R. and Wilson,R.
  Public Soybean EST Project
  Unpublished (1999)
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.edu
  This clone is available through: Resgen, Invitrogen Corp. 2130
  South Memorial Parkway Huntersville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: cc@resgen.com
  Insert length: 1779      Std Error: 0.00
  High quality sequence stop: 414.
FEATURES
  source
  1..588
  /organism="Glycine max"
  /db_xref="taxon:3847"
  /clone="GENOME SYSTEMS CLONE ID: Gm-c1067-1734"
  /clone.lib="Gm-c1067"
  /tissue.type="germinating shoot, 3 day old seedling, auxin
  treatment"
  /lab_host="DH10B"
  /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
  XhoI; The cDNA library was constructed from mRNA isolated
  from germinating shoots of 3 day old seedling for the
  cultivar Williams 82. The seedlings were germinated in a
  growth chamber using germination paper in a solution
  containing 100ppm auxin. Complementary DNA was synthesized
  from mRNA using a primer consisting of a poly(dT) sequence
  with a XhoI restriction site. EcoRI adapters were ligated
  to the blunt-ended cDNA fragments followed by XhoI
  digestion. The cDNA fragments were directionally cloned
  into the EcoRI-XhoI restriction site of the pBluescript
  vector. The ligated cDNA fragments were transformed into
  DH10B host cells (GibcoBRL). This library was constructed
  in the laboratory of Dr. Randy Shoemaker."
BASE COUNT      182 a      114 c      124 g      168 t
ORIGIN
Alignment Scores:
Pred. No.:      8.81e-117      Length:      588
Score:          1042.00      Matches:      194
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.49%      Mismatches: 0
Query Match:    47.65%      Indels:      0

```

DB: 12 Gaps: 0
US-09-720-451-6 (1-408) x BF071278 (1-588)

OY 3 ValIleValIThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsnAspCys 22
:::|||||
DB 4 CTGATGTGCACCAATTTTATGATAGAGATATATTCCTCAATTTGATGACGACTGT 63
OY 23 ArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsnTyrLys 42
|||||
DB 64 CCCCAAATTGGATATACCTGTCTTATTTACCTGGAATTATGCCCATTAATAATTACAG 123
OY 43 GlyPheIleArgMetCysGlyPheCysLysThrLysIleProAlaAspIleMetAlaAla 62
|||||
DB 124 GGCTTATCCGATGACTGATGATTTGCAAAACAAAGATCACGCTGACATTATGCTGCT 183
OY 63 LeuGluProIleLysAspAsnGlnGluAlaValLysAlaTyrGlyIleHisLeuGlyThr 82
|||||
DB 184 TTAGAGCCATTCAGACATGAGAGCTGCAAGGCTTATGGAATTCACCTGGACT 243
OY 83 GluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsn 102
|||||
DB 244 GAAATGTGCAAAAGATTTTACCTCATGTGAATTAGACATTGCATCTTATACATAAT 303
OY 103 MetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSerLysVal 122
|||||
DB 304 ATGGAGAAATCTGCATTTGCATTAATAAGACCTTGCCCTAATTGAAGAGTCCAAAGTT 363
OY 123 SerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGluAspValArg 142
|||||
DB 364 TCTAGTCTTACCTTGGAGAGCCCTGCAAATGTTTCCGTGTAAGAGATGCCGT 423
OY 143 ProIlePheTyrPalaAsnArgProLysSerTyrIleSerArgThrIleGlyTyrAspGln 162
|||||
DB 424 CCAATCTTTTGGCAAAATGCACCAAAAGCTACATATCAAGGACCATAGATGGGATCAA 483
OY 163 TyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyr 182
|||||
DB 484 TACCCACATGGGCGGTGGGTGATTCCTGTATCCATCATATGATGTCATATCTGATTAT 543
OY 183 GlnPheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
|||||
DB 544 CAGTTTCATCGGCGCACGTCACGGGACAGAAAGCTTGTTGAAGAA 588

Search completed: February 15, 2003, 00:51:19
Job time : 2170 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 14, 2003, 21:53:11 ; Search time 273 Seconds

(without alignments)
3365.624 Million cell updates/sec

Title: US-09-720-451-6

Sequence: 1 ARVIVTQLFYDTDFILKEVN.....VSLVNDYINGDLFAVFAADF 408

Scoring table:

| | | |
|-------------|-------------|-------------|
| BLOSUM62 | Xgapop 10.0 | Xgapext 0.5 |
| Ygapop 10.0 | Ygapext 0.5 | |
| Fgapop 6.0 | Fgapext 7.0 | |
| Delop 6.0 | Delext 7.0 | |

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-O=/cgn2.1/USPTO.spool/US09720451/runat_05022003_074809_2374/app_query.fasta.1.563
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cd1
-LIST=45 -DOCCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720451.ecgn.1.1.0.@runat.05022003.074809_2374 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEROUDRY -NEG_SCORES=0 -MAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
N_Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|----------------------|
| 1 | 1834 | 83.9 | 1821 | 21 | AAC43311 | Arabidopsis thaliana |
| 2 | 1834 | 83.9 | 1942 | 21 | AAC49168 | Arabidopsis thaliana |
| 3 | 1760 | 80.5 | 1335 | 21 | AAC36967 | Arabidopsis thaliana |
| 4 | 1184 | 54.1 | 1493 | 21 | AAC44829 | Arabidopsis thaliana |
| 5 | 998 | 45.6 | 558 | 21 | AA550061 | Soybean 5,10-methyl |
| 6 | 979 | 44.8 | 2187 | 22 | AAC91206 | Human methylenele |
| 7 | 979 | 44.8 | 2219 | 17 | AAO96889 | Human methylenele-te |
| 8 | 979 | 44.8 | 2219 | 22 | AAD20463 | Human methylenele |
| 9 | 979 | 44.8 | 2230 | 17 | AAO96949 | Human methylenele-te |
| 10 | 979 | 44.8 | 2230 | 21 | AA550633 | Human methylenele |
| 11 | 979 | 44.8 | 2230 | 22 | AAD20462 | Human methylenele |
| 12 | 979 | 44.8 | 2230 | 24 | ABK13501 | DNA encoding human |
| 13 | 979 | 44.8 | 2230 | 24 | ABK13539 | Human methylenele |
| 14 | 979 | 44.8 | 2230 | 24 | ABK13540 | Human methylenele |
| 15 | 979 | 44.8 | 2230 | 24 | ABK13541 | Human methylenele |
| 16 | 979 | 44.8 | 2230 | 24 | ABK13545 | Human methylenele |
| 17 | 979 | 44.8 | 2230 | 24 | ABK13546 | Human methylenele |
| 18 | 979 | 44.8 | 2230 | 24 | ABK13550 | Human methylenele |
| 19 | 978 | 44.7 | 2230 | 24 | ABK13542 | Human methylenele |
| 20 | 976 | 44.6 | 2230 | 24 | ABK13547 | Human methylenele |
| 21 | 973 | 44.5 | 2230 | 24 | ABK13543 | Human methylenele |
| 22 | 971 | 44.4 | 2230 | 24 | ABK13548 | Human methylenele |
| 23 | 969 | 44.3 | 2230 | 24 | ABK13544 | Human methylenele |
| 24 | 950 | 43.4 | 832 | 24 | ABN98922 | Arabidopsis thaliana |
| 25 | 742 | 33.9 | 3283 | 21 | AA550634 | Human methylenele |
| 26 | 706.5 | 32.3 | 3181 | 21 | AA550635 | Mouse methylenele |
| 27 | 415 | 19.0 | 23748 | 23 | AAO55229 | Human reproductive |
| 28 | 415 | 19.0 | 23748 | 23 | AB198112 | Human testicular a |
| 29 | 386.5 | 17.7 | 451 | 21 | AA550062 | Wheat 5,10-methyl |
| 30 | 255 | 11.7 | 159 | 24 | AB175279 | Corn tassell-derive |
| 31 | 188 | 8.6 | 2440 | 23 | AA593090 | DNA encoding novel |
| 32 | 187.5 | 8.6 | 349980 | 21 | AA597269 | Neisseria meningit |
| 33 | 182 | 8.3 | 876 | 24 | AA597261 | Neisseria meningit |
| 34 | 172.5 | 7.9 | 283 | 22 | AAD20482 | Human MTRFR gene e |
| 35 | 172.5 | 7.9 | 283 | 22 | ABK13510 | Human methylenele |
| 36 | 169 | 7.7 | 235 | 22 | AAD20480 | Human MTRFR gene e |
| 37 | 169 | 7.7 | 235 | 24 | ABK13508 | Human methylenele |
| 38 | 168 | 7.7 | 640681 | 24 | ABA92787 | Buchnera sp. genom |
| 39 | 166 | 7.6 | 235 | 22 | AAD20491 | Mouse MTRFR gene e |
| 40 | 166 | 7.6 | 235 | 24 | ABK13519 | Mouse methylenele |
| 41 | 160.5 | 7.3 | 283 | 22 | AAD20493 | Mouse MTRFR gene e |
| 42 | 160.5 | 7.3 | 283 | 24 | ABK13521 | Mouse methylenele |
| 43 | 158 | 7.2 | 978 | 22 | AAH67344 | C glutamicum codin |
| 44 | 158 | 7.2 | 1005 | 22 | AAH72089 | Corynebacterium gl |
| 45 | 158 | 7.2 | 349980 | 22 | AAH68531 | C glutamicum codin |

ALIGNMENTS

RESULT 1
AAC43311 standard; DNA; 1821 Bp.

AAC43311:

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 38803.

Hybridisation assay; genetic mapping; gene expression control;

protein identification; signal transduction pathway;

metabolic pathway; promoter; termination sequence; ss.

Arabidopsis thaliana.

EPI033405-A2.

PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130691.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143547.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147735.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0153659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

| Pred. No.: | 1,55e-179 | Length: | 1821 |
|------------------------|-----------|---------------|------|
| Score: | 1834.00 | Matches: | 334 |
| Percent Similarity: | 91.15% | Conservative: | 37 |
| Best Local Similarity: | 82.06% | Mismatches: | 36 |
| Query Match: | 83.86% | Indels: | 0 |
| DB: | 21 | Gaps: | 0 |

US-09-720-451-6 (1-408) x AAC43311 (1-1821)

```
OY 1 AIAATGValIleValThGlnLeuPheTyRAspThrAspIlePheLeuLysPheValAsn 20
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 595 GGTGATCTGATGTGTAATCAAGCTTTCTATGATGATGATATATCTCAAGTTGTGAT 654

OY 21 AspCysArgGlnIleGlyLThrCysProIleValProGlyIleMetProIleAsnAsn 40
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 655 GATTGGCGCAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 714

OY 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaAspIleMet 60
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 715 TACAGAGGCTTTTGGGTAAGTACTGTTCTCAAGACTCAAGATCAAGTGAAGGTGATG 774

OY 61 AIAAlaLeuGluProIleLysAspAsnGluAlaValIleValAlaTyGlyLThrIleIleu 80
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 775 GCTGCTTGAGCTATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 834

OY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGlyLLeuLysThrLeuHisLeuTyrThr 100
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 835 GGAACAGAGATGTGTAAAAAGATGTGGCTCATGGAGTCAAGTCTCTCATCTCTACACA 894

OY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluGluSer 120
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 895 TTGAACATGAGAAATCTGCTTGCAATATGATGAATCTTGATGATGATGATGATGATG 954
```

```
OY 121 LysValSerArgSerLeuProThrParGArgProAlaAsnValPheArgValLysGluAsp 140
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 955 AAAATTTTCGCTTCTTTACCTCGGAGACGCCCTGCAATGTTTTCGACTACAAAGAAAT 1014

OY 141 ValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGlyTyrP 160
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1015 GTGGCGCCCATTTTCTGGGCAAAAGCTCCAAAGACTACATTTCTAGAACCAAGGCGTGG 1074

OY 161 AspGlnTyrProHisGlyLysArgTyrPglLysAspSerCysAsnProSerTyrGlyAlaLeuSer 180
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1075 GAAGATTTCCACCAAGCGCGGGGGGTGATTCAGCAGAGCTTCATATAGGTGACTCTCG 1134

OY 181 AspTyrGluPheMetArgProArgProAlaArgAspLysLysLeuValGluGluTyrPalaVal 200
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1135 GATCATCAGTTCTTCACGCTCCGAGACCTGCACAAAGAACCTTCACAAAGAAATGGTGTGC 1194

OY 201 ProLeuLysSerValGluLysAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLysIleu 220
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1195 CCACGTAAAGGTGTTCAAGATATTCAGAGAAATTCAGAGACCTGCTGCAAAACCTT 1254

OY 221 ArgSerAsnProTyrPserGluLeuAspGlyLeuGlnProGluThrLysIleIleAsnGlu 240
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1255 AAAAGCAAGTCCCTGCTCAATTAAGTAGGACTCCACCCAGACACAAAGATTAATAACGAG 1314

OY 241 GlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnProAlaValAlaAsn 260
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1315 CAACGTATAAAGTCAACACCAAGGCTTGTGACCATCAATAGCCAAACACACAGTCAAC 1374

OY 261 GlyLysLysSerAspSerProThrValGlyTyrPglGlyProGlyGlyTyrValTyrGln 280
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1375 GCCAGAGATCTGATGCCCAACACTGTTGGATGGGAGGTCTGTGGAAATATATACCAA 1434

OY 281 LysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLysGluIleIle 300
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1435 AAGGCTTATCTAGAGTCTTCTCTCAAGAGCAAAATTAAGTACAGTGTGGAGAAATGC 1494

OY 301 LysAspArgThrSerLeuThrTyrMetAlaValAlaAsnLysAspLysSerTyrLysSerAsn 320
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1495 AAGGCTTGGCATCGATCTTACTTACATGGCTGTGAACAAAGAGCAACAGTGGTATGAAC 1554

OY 321 ValGlyGlnThrAspValAlaAsnAlaValThrTyrGlyValPheProAlaLysGluIleIle 340
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1555 ACTGCCCAACCTGATGGAATGCTGTAACTTGGGAGTTCCTCCGGCTAAAGAAATCAAT 1614

OY 341 GlnProThrIleValAspProValSerPheAsnValTyrLysAspGluAlaPheGluIle 360
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1615 CAACCAACCATTTGTCATCCCGCTAGCTTCAAGTCTGGAAGATGAAGATTTGAGACT 1674

OY 361 TrpSerArgGlyTyrPalaSerLeuTyrProGluAspGluAlaSerArgLysLeuValGlu 380
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1675 TGGTCAAGAGAGTGGGCTACTGTACCCAGAGCCGACCTTCAGAAACTTGCTCGAG 1734

OY 381 GluValGlyGlySerHisPheLeuValSerLeuValAspAsnAspTyrIleAsnGlyAsp 400
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1735 GAGGTGAAGAACACTCTATTTGGTGAAGCTGTGTGAGAGAACATTCATCATGTGTGAC 1794

OY 401 LeuPheAlaValPheAlaAsp 407
   ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1795 ATATTCCGGGTGTGTGCTGAT 1815

RESULT 2
AAC49168
ID AAC49168 standard; DNA; 1942 BP.
XX
AC AAC49168:
XX
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 60181.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
```

XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125288.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130649.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134570.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

```
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

```
Pred. No.: 1.69e-179 Length: 1942
Score: 1834.00 Matches: 334
Percent Similarity: 91.15% Conservative: 37
Best Local Similarity: 82.06% Mismatches: 36
Query Match: 83.86% Indels: 0
DB: Gaps: 21
```

US-09-720-451-6 (1-408) x AAC49168 (1-1942)

```
QY 1 ALAARYAIIIEVALIITHRGInLeuPheTYRASPThrAspIlePheLeuLysPheValAsn 20
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 559 GCTGATCGATGATTACTACACGCTTTTCATGATACTGATATATATCTCAAGATTGTGAAT 618
QY 21 ASPCYASRGInIIEGLIIEThRCysProIleValProGLIIEMetProIleAsnAsn 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 GATTGTCGGCAAAATGGGATTAGTTGCCATTGTCTCGAATTAATGACCTATTAAATAC 678
QY 41 TYRISGLYPheIleArqMetThGLYPheCysLysThrLysIleProAlaAspIleMet 60
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 679 TACAGAGCGCTTTTGGCGATGATCGGTTTCGCAAGACTAAGATACCAAGTTGAGGTATG 738
QY 61 ALAAlaLeuGIUProIleLysASPasnGLIuAlaValLysAlaTYRGLIIEHisLeu 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 GCTGCTTGAGCCTATCAACAGATACGAAGAACTGTGAAGACCTATGATTCACCTT 798
QY 81 GLYThrGluMetCysLysLysIleLeuAlaHisGLIYIleLysThrLeuHisLysLeuTYR 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 GGAACAGAGATGTATAAAAGATGTGGCTCATGAGAGTCAAGTCTCTTCATCTCTACACA 858
```

```
QY 101 LeuAsnMetClyuysSerAlaLeuAlaIleLeuMetAsnLeuGLIueuIleGLIueuIser 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 TTGAACATGGAGAAATCTGCTCTTGCAATATGATGAATCTGTGATGATGATGATGATG 918
QY 121 LYSValSerArqSerLeuProTrpArqArqProAlaAsnValPheArqValLysGLIuAsp 140
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 919 AAAATTTCGTCTTTTACCCTGGAGACGCCCTGCAGAAAGTTTTCGTCCTAGGAGAGAT 978
QY 141 ValArqProIlePheTrpAlaAsnArqProLysSerTYRILESerArqThrIleGLIuTrp 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 GTGCGCCCATTTTCTGGGCAAAACGTCACAAAGAGCTACATTCTTCAAGAACAGGCGTGG 1038
QY 161 ASPGLINTYRProHisGLIYArqTrpGLIYAspSerCysAsnProSerTYRGLYAlaLeuSer 180
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1039 GAAGATTTCACACAGGCGCGGTGGATTCACGCGAGCTTCATATGCTGACACTCTCG 1098
QY 181 ASPTYRGLNPheMetArqProArqAlaArqAspLysLysLeuValGLIuTrpAlaVal 200
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1099 GATCATCAGTTCTCAGCTCCGCGAGCAGCTGACACAGAACCTTCACACAAATGGCTGTTC 1158
QY 201 PROLeuLysSerValGLIuAspIleTYRGLIuArqPheArqLeuTYRCysLeuGLIYysLeu 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1159 CCAGTGAAGAGTGTGAAGATATTCAGAGAAATTCAGAGAGCTGCGCTTGAAAACCTT 1218
QY 221 ARGSerAsnProTrpSerGLIueuAspGLIueuInProGLIuThrLysIleIleAsnGLIu 240
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1219 AAAAGCATGCTCGTGTGATAGATAGATGATCCACCGAGACAGATTAATTAACGGG 1278
QY 241 GLNLeuGLIuLysIleAsnThrLysGLIYpHeLeuThrIleAsnSerGLNProAlaValAsn 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1279 CAACGTATAAAGTCAACATCCCAAGGCTTCTGACCATCAATAAAGCAACATCAGTCAAC 1338
QY 261 GLYGLIuLysSerAspSerProThrValGLIYTrpGLIYProGLIYGLIYTYRValTYRGLN 280
    ||| ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1339 GCCAGAGAGATCGATTCCCAACTGTTGATGGGAGGCTCGTTGATATATATACCA 1398
QY 281 LYSAlaTYRValGLIuPheCysSerLysGLIuLysLeuAspAlaLeuValAspLysCys 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1399 AAGGCTTATCTAGAGTTCTTCTGCTCAAGAGAAATTAATGACAGTGGTGAAGAAATTC 1458
QY 301 LYSASPArqThrSerLeuThrTYRMetaLalaValAsnLysAspGLYSerTrpLysSerAsn 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1459 AAAGCTTTGGCATCGATTACTTACATGGCTGTGAACAAAGAGAACAGTGTATTCGAAC 1518
QY 321 VALGLYGLNThrAspValAsnAlaValThrTrpGLYValPheProAlaLysLutIle 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1519 ACTGCCCAAGCTGATGTGAATGCTGTAACTGGGAGTTTCCCGCTTAAGAAATTCATT 1578
QY 341 GLNProThrIleValAspProValSerPheAsnValTrpLysAspGLuAlaPheGLIuIle 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1579 CAACCAACCATTTGTCATCCGCTGACCTTCACAGCTTGGAAGAGATGACATTTGAGACT 1638
QY 361 TRPSetArqGLYTrpAlaSerLeuTYRProGLIuAspGLuAlaSerArqLysLeuValGLIu 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1639 TGGTCAAGAAAGCTGAGCTTACTGTATCCAGAAAGCCGACCTTCCAGAAATCTTGCGAG 1698
QY 381 GLUValGLIYGLYSerHisPheLeuValSerLeuValAspAsnAspTYRILEAsnGLIYAsp 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1699 GAGGTGAAGAACGCTACATATTGTGTAAAGCTGTGTGAAGAACGATTACATCAATGTGTAC 1758
QY 401 LeuPheAlaValPheAlaAsp 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1759 ATATTGCGGCTGTTCCTGAT 1779
```

RESULT 3

AAC36967 standard; DNA; 1335 BP.

AAC36967;

17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15706.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142035.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

| Pred. No.: | 4,48e-172 | Length: | 1335 |
|------------------------|-----------|---------------|------|
| Score: | 1760.00 | Matches: | 320 |
| Percent Similarity: | 90.77% | Conservative: | 34 |
| Best Local Similarity: | 82.05% | Mismatches: | 36 |
| Query Match: | 80.48% | Indels: | 0 |
| DB: | 21 | Gaps: | 0 |

US-09-720-451-6 (1-408) x AAC36967 (1-1335)

QY 18 PheValAsnAspCysArgGlnIleGlyIleThCysProIleValProGlyIleMetPro 37
DB 1 TTTGTGAAATGATGTCGCGAAATGGATTAATTTGCCATTTGCTCGTAATTAAGCCT 60
QY 38 IleAsnAspTyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAla 57
DB 61 ATTAATTAACACTACAGAGGCTTCTTGCGATGATGCTTTCTGCAAGACTAAGATACAGATT 120
QY 58 AspIleMetAlaAlaLeuGluProIleLysAspAsnGluGluAlaValLysAlaTyrGly 77
DB 121 GAGGTATGCTGCTGCTTGAGGCTATCAAGCATACAGAAACCTGTGAAGCCTATGCT 180
QY 78 IleHisLeuGlyThrGluMetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHis 97
|||||

DB 181 ATTCACTTGAAACAGAGATGTTGTAATAAAGATGTGGCTCATGAGTCAAGTCTTCAT 240
QY 98 LeuTyrThrLeuAsnMetCyluysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIle 117
DB 241 CTTTACACATTAACATGAGAAATCTGCTTTGCAAAATTAATGATCAATCTTGATGATT 300
QY 118 GluGluSerLysValSerArgSerLeuProTrrPAArgArgProAlaAsnValPheArgVal 137
DB 301 GATGAGTCCAAAATTTCTGCTTTTACCTTGAGACGCCCTGCAAAAGTTTCCGTACT 360
QY 138 LysGluAspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThr 157
DB 361 AAGGAAGATGCGGCCCATTTCTGCGCAACCGTCCAAAGAGTACATTCTGAAGCC 420
QY 158 IleGlyTrrAspGlnTyrProHisGlyArgTrrPglyAspSerCysAsnProSerTyrGly 177
DB 421 AAGGCTGGGAAGACTTCCCAAGCGGTGGGTGATTCACGACGCTTCATATGCT 480
QY 178 AlaLeuSerAspTyrGlnPheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
DB 481 GCACCTCGCATCATCAGTTCTCAGCTCCGCGACGACGACGACAAAGCTTCACACAGA 540
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
DB 541 TGGGTGTCCCACTGAAAGTGTGAAGATTTCAAGAAATTCAGAGAAATTCAGAGACTGCTT 600
QY 218 GlyLysLeuArgSerAsnProTrrSerGluLeuAspGlyLeuGlnProGlyIleThrLysIle 237
DB 601 GGAACCTTAAGAGCGCTGCTGTAATTAAGATGACCTCCACGACGACGACAAAGATA 660
QY 238 IleAsnGluGlnLeuGluLysIleAsnThrLysGlyPheLeuThrIleAsnSerGlnPro 257
DB 661 ATAAAGACACACTGGTAAAGTCAACTCAAGAGCTTCTTGACCATCAATACCAACCA 720
QY 258 AlaValAsnGlyGluLysSerAspSerProThrValGlyTrrPglyLysProGlyLysTyr 277
DB 721 TCAGTCAACGCCGAGATGATTCCTCCCAACTGTGGATGGGAGGTCCTGTGATAT 780
QY 278 ValTyrGlnLysAlaTyrValGluPhePheCysSerLysGluLysLeuAspAlaLeuVal 297
DB 781 GTRATCCAAAGAGCTTATCTAGATTCCTGCTCAAGGAGAAATTAAGATGCACTGGTG 840
QY 298 AspLysCysLysAspArgThrSerLeuThrTrrMetAlaValAsnLysAspGlySerTrr 317
DB 841 GAGAAATGCAAGCTTTGCCATCGATTTCAATGAGGCTGTCACAAAGAGACAGTGG 900
QY 318 LysSerAsnValGlyGlnThrAspValAsnAlaValIleTrrPglyValPheProAlaLys 337
DB 901 GTATCGAACACTGTCCAAGCTGATGTAATGCTGTAACCTGGGAGTTCCTCCGCTAAG 960
QY 338 GluIleIleGlnProThrIleValAspProValSerPheAsnValTrrLysAspGluAla 357
DB 961 GAATCATTTCAACCAACTTGTGATCCCGCTACCTCAACGTCGTGGAAGAGTAACA 1020
QY 358 PheGluIleTrrSerArgGlyTrrPalaSerLeuTrrProGluAspGluAlaSerArgLys 377
DB 1021 TTGAGACTTGTCTCAAGAAAGCTGGCTAATCTGTACCCAGAAAGCCGACCTTCCAGAAAC 1080
QY 378 LeuValGluGluValGlyLysThrHisPheLeuValSerLeuValAspAsnAspTyrIle 397
DB 1081 TTGCTCGAGAGAGTGAAGACAGCTACTATTGTTGAACCTTGTTGAGAACGATTACATC 1140
QY 398 AsnGlyAspLeuPheAlaValPheAlaAsp 407
DB 1141 AATGGGACATATTCGCGTCTTTGCTGAT 1170
RESULT 4
AAC44829
ID AAC44829 standard; DNA; 1493 BP.
XX AAC44829;
AC
XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 44287.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PM EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0133484.
PR 05-MAY-1999; 99US-0133485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144682.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147028.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149176.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.


```
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1.89e-112 | Length: | 1493 |
| Score: | 1194.00 | Matches: | 213 |
| Percent Similarity: | 89.26% | Conservative: | 28 |
| Best Local Similarity: | 78.89% | Mismatches: | 29 |
| Query Match: | 54.14% | Indels: | 0 |
| DB: | 21 | Gaps: | 0 |

US-09-720-451-6 (1-408) x AAC44829 (1-1493)

```
OY 1 ALaarYValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
DB 642 GCACATTTGATGTGACACACCTTTCTATGATGATATATTCCTCAAGTTTGTAAT 701
OY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
DB 702 GATTGTGCGCAATCGGATTAATGTGCTTCCTGGAATTAATGCTATTTCACAC 761
OY 41 TyrIysGlyPheIleArgMetThrGlyPheCysIysThrIysIleProAlaAspIleMet 60
DB 762 TACAAGGGGTTCTTGCGATATGCGTGTTTCGTAAAGACCAAGATACCCGCTGAGCTACT 821
OY 61 AlaAlaLeuGluProIleLysAspAsnGluAlaValLysAlaTyrGlyIleHisLeu 80
```

```
DB 822 GCTGCCTTAGACCATTAATGATATGACGAGGCTGTAAAGCCTATGGAATTCACCTT 881
OY 81 GlyThrGluMetCysIysIleLeuAlaHisGlyIleLeuThrLeuHisLeuYThr 100
DB 882 GCACAGAAATGTGCAAAAGATTTGGCCATGATATCCTTCCTATCTTACACA 941
OY 101 LeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluLys 120
DB 942 TTGAACGTGCAACAATCASCATTTGGATATTAATCAACCTTGCTGATATGATCA 1001
OY 121 LysValSerArgSerLeuProIlePargArgProAlaAsnValPheArgValLysGlu 140
DB 1002 AAATTTCTGCTCTCTACCTTGAGACGCCCTGCAAAATGTTTCGTTACTAAGAA 1061
OY 141 ValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly 160
DB 1062 GTTCGCCCAATTTTCGGCCAAACCGTCCAAAGAGCTATATCTGAACAAAGGCTGG 1121
OY 161 AspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 180
DB 1122 AATGACTTCCACATGACGCTTGCGGTATTCACACAGTCGACATACGATACCTTGC 1181
OY 181 AspTyrGlnPheMetArgProAlaArgAspLysLysLeuValGluIleTrpAlaVal 200
DB 1182 GATTATCAGTTTCGGCCCAAAAGACGTGACAAAGACCTTCACAAAGATGGTCTGC 1241
OY 201 ProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeuGlyLys 220
DB 1242 CCACGTAAAGCATTAAGATGTTCAAGAGAAATTCAAAGACCTCTGCATTTGAA 1301
OY 221 ArgSerAsnProTrpSerGlnLeuAspGlyLeuGlnProGlyIleThrLysIleLeu 240
DB 1302 AAAAGCAGCCCATGCTCTAATTAATGATGACTCCAGCCAGACAAAGATCTAAATG 1361
OY 241 GlnLeuGluLysIleAsnThrIysGlyPheLeuThrIleAsnSerGlnProAlaVal 260
DB 1362 CAACCTGGAAATCAACGCAACGCTTCCTGACCATTAATAGCAACATCAGTCAAT 1421
OY 261 GlyGluLysSerAspSerProThrValGly 270
DB 1422 GCAGCCAAATCCGATTCCTCAGCTATTGCT 1451

RESULT 5
AAZ50061
ID AAZ50061 standard; cDNA; 558 BP.
XX
AC AAZ50061;
XX
DT 04-MAY-2000 (first entry)
XX
DE Soybean 5,10-methylenetetrahydrofolate reductase cDNA clone.
XX
KW 5,10-methylenetetrahydrofolate reductase; MTHFR; soybean; inhibitor;
methionine synthesis; tetrahydrofolate metabolism enzyme; herbicide; ss.
XX
OS Glycine max.
XX
FH
FT key Location/Qualifiers
FT mat_peptide 3..557
FT /*tag_a
FT /partial
FT /product= "5,10-methylenetetrahydrofolate reductase"
FT /note= "encodes amino acids Ile4-Arg188 of the enzyme"
FT /EC_number= "1.7.99.5"
XX
PN WO200004163-A1.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15916.
XX
PR 15-JUL-1998; 98US-0092869.
```


| | | | |
|----|------|--|-------|
| OY | 61 | AlAlAlLeuGluProIleuLysAspAsnGluAlaValAlaValAlaValGlyGlyIleHisLeu | 80 |
| Db | 847 | GACGTGTAGAGCCAAATCAAAAGACAAGATGCTGCATCCGAACTATGACATGACGTG | 906 |
| OY | 81 | GlyThrLeuMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr | 99 |
| Db | 907 | GGCGTAGCGCTGTGCAGAGACCTTCTGGCGAGTGGCTTGGTCCAGAGGCTGCACCTTGC | 966 |
| OY | 100 | ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu | 119 |
| Db | 967 | ACCCCTCAACCGCGAGAGGTCTACACAGAGAGTGTCTGAAGCGGCTGGGATGTGACATG | 1026 |
| OY | 120 | SerLysValSerArgSerLeuProTyrPArgArgProAlaAsnValPheArgValGlu | 139 |
| Db | 1027 | GACCCC---AGGGCTCCCTTACCTTGGCTGTACGTGCCCCACCCCAAGCGCGAGAGGA | 10833 |
| OY | 140 | AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly | 159 |
| Db | 1084 | GATTAAGTCCCATCTTCTGGCGCTCCAGACCAAAAGATTACATCTACCGTACCGACGAG | 1143 |
| OY | 160 | TrpAspGlnTyrProHisGlyArgTyrPrgLysSerCysAsnProSerTyrGlyAlaLeu | 179 |
| Db | 1144 | TGGAGCAGATTCCTTAACGGCCCTGGGGAATTCCTTCCCTGCGCTTGGGAGGTG | 12030 |
| OY | 180 | SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu | 197 |
| Db | 1204 | AAGGACTACTACGCTCTTCTACGTGAAGAGCAAGTCCCCCAAGAGAGAGCTGTGTAAAGTG | 1265 |
| OY | 198 | TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu | 217 |
| Db | 1264 | TGGGGGAGGAGCGTACGACATGAAGCAAGTGTCTTGAAGTGTCTTGTCTTACCTCTCG | 13233 |
| OY | 218 | GlyLysLeuArgSerAsn-----ProTyrSerGluLeuAspGly | 230 |
| Db | 1334 | GGAGACCAAAACCGGAATGTGTCAAAAGTACTGTGCTGCTGCCTGGAGACAT---GAGCCC | 1380 |
| OY | 231 | LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe | 250 |
| Db | 1381 | CTGGCGGCTGTAGACCAACCTGTGAAGAGAGACTCTCGGGTGAACCGCACGAGGCATC | 1440 |
| OY | 251 | LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly | 270 |
| Db | 1441 | CTCACCATCACTACGACCCCAATCAATCAAGGGGAACCGCTCCGACCCCACTGTGGAC | 1500 |
| OY | 271 | TyrGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys | 290 |
| Db | 1501 | TGGGGCCCAACGGGGGCTATGCTTCTTCCAAAGACGCTACTATGAGTTTTCACCTCCGC | 1560 |
| OY | 291 | GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet | 309 |
| Db | 1561 | GAGACACGGGAAGCACTTGTCAAGTCTGAAAGATCAAGTACGAGCTCCGGGTATTAATCAC | 1620 |
| OY | 310 | AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAsnAlaVal | 329 |
| Db | 1621 | CTTGTCATGTGAAGGGTGAAGAATCATCAACCAATGCCCTGAACTGACGCGAATCTGTC | 1680 |
| OY | 330 | ThrTyrGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer | 349 |
| Db | 1681 | ACTTGGGGCATCTTCCCTGTGGCGAGAGATCATCAACCCACCGTATGTGATCCCTGACG | 1740 |
| OY | 350 | PheAsnValTyrPlysAspGlyAlaPheGluIleTyrSerArgGlyTrpAlaSerLeuTyr | 369 |
| Db | 1741 | TTTCATGTTCTGGAAGACAGAGGCTTTGGCCCTGTGATTGAGCGGTGGGAAAAGCTGAT | 1800 |
| OY | 370 | ProGluAspGlnAlaSerArgLysLeuValGluGluValAlaGlyLysSerHisLeuVal | 389 |
| Db | 1801 | GAGGAGGAGTCCCGCTCCGACCATCATCAAGACATCCACGACAACTACTTCCGTGTC | 1860 |
| OY | 390 | SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp | 407 |
| Db | 1861 | AACCTGTGTGGACATGATCTTCCATCTGCAACTGCTGCTTGGCGAGGTGTGGAGAAC | 1917 |

| | |
|----------|---|
| RESULT | 7 |
| AAT09689 | |
| ID | AAT09689 standard; cDNA; 2219 BP. |
| XX | |
| AC | |
| XX | AAT09689; |
| XX | |
| DT | 15-OCT-1996 (first entry) |
| XX | |
| DE | Human methylene-tetrahydrofolate-reductase cDNA. |
| XX | |
| XX | Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy; |
| KW | cardiovascular disease; neurological disease; folic acid metabolism; |
| KM | EC-1.5.1.20; enzyme; SS. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | |
| FT | Key location/Qualifiers |
| CDS | 13..1980 |
| FT | /*tag= a |
| PX | |
| PN | W09533054-A1. |
| PD | 07-DEC-1995. |
| XX | |
| PF | 25-MAY-1995; 95WO-CA00314. |
| PR | 26-MAY-1994; 94GB-0010620. |
| XX | |
| PA | (UYMC-) UNIV MCGILL. |
| XX | |
| PI | Goyette P, Rozen R; |
| XX | |
| DR | WPJ; 1996-030565/03. |
| DR | P-PSDB; AAR88358. |
| XX | |
| PT | Human methylene:tetra:hydro:folate reductase cDNA probe - for |
| PT | detection of sequence abnormalities in methylene:tetra:hydro:folate |
| PT | metabolism disorders |
| XX | |
| PS | Disclosure; Fig. 6A-6C; 66pp; English. |
| XX | |
| CC | This sequence encoding human MTHFR has been localised to chromosome |
| CC | 1p36.3. It may be used for the construction of DNA probes which may |
| CC | be used for the identification of sequence abnormalities in patients |
| CC | with severe or mild MTHFR deficiency. The resulting probe may also |
| CC | be used in gene therapy to produce the MTHFR protein. |
| XX | |
| SQ | Sequence 2219 BP; 501 A; 656 C; 620 G; 442 T; 0 other; |
| | |
| | Alignment Scores: |
| | |
| | Pred. No.: 5,19e-91 Length: 2219 |
| | Score: 979.00 Matches: 188 |
| | Percent Similarity: 63.72% Conservative: 79 |
| | Best Local Similarity: 44.87% Mismatches: 138 |
| | Query Match: 44.76% Indels: 14 |
| | DB: 17 Gaps: 7 |
| | |
| | US-09-720-451-6 (1-408) x AAT09689 (1-2219) |
| OY | 1 AAlaArGValIleValtThrGlInLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20 ::: :: :: :: :: :: :: :: :: :: :: :: :: Db 676 GCCGATTTCATCATCATCACGCACCTTTCTTGAGGCTGACACAACTTCCTCCGCTTGTGAAG 735 OY 21 AsPySaArGInIleGlyIleThrCySproIleValProGlyIleMetProIleAsnAsn 40 ::: :: :: :: :: :: :: :: :: :: :: :: :: Db 736 GCATCCACCAGCATGGGACTTACTGCCCAACGTCTCCCGGGATCTTCCATCCAGGCG 795 OY 41 TrLyLSgLYPheIIeaRGeMeTrhGlyYPheCysLTsrThySLiePrOLAAspIleMet 60 ::: :: :: :: :: :: :: :: :: :: :: :: :: Db 796 TRCAcAcTCCCTTCGGCAGCTTGTGAAGTGTCCAAGCTGGAGGTCGACACAGAGATGAAG 855 OY 61 AlAlaIleUGInUProIleLysAspAsnGLInuAlaValLysAlaTyrgLyILenHisLeu 80 |

```
Db      856 GACGTATTGAGCAATCAAGACAGATGCTGCCATCGCAACTATGATGCAGCTG 915
      81 G1YThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTy 99
      916 GCCGTAGAGCTGTGCCAGAGGCTTGTGGCCAGTGGCTTGGCCAGGCTTCCTTCTAC 975
      100 ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuGlu 119
      976 ACCCTCAACCGGAGATGGCTACACAGAGTGCTCAACCGGCTGGGAGATGGAGCTGAG 1035
      120 SerLysValSerArgSerLeuProTyrParArgProAlaAsnValPheArgValLysGlu 139
      1036 GACCCC---AGGGTCCCTACCTGCGCTCTCAGTGGCCACCCCAAGGCGGAGAGGAA 1092
      140 AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly 159
      1093 GATGTACGTCCCATCTTCTGGGCTCCAGACCAAGAGTTACATCTACCGTACCCAGAG 1152
      160 TrpAspGlnTyrProHisGlyArgTrpGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
      1153 TGGGAGAGATGCCCTACAGCGCGCTGGGCAATCTCTCCCTCCCTTGGGAGCTG 1212
      180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu 197
      1213 AAGGACTACTACCTCTTCTACCTGAGAGACAACTCCCCCAAGAGAGAGCTCTGAAGATG 1272
      198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu 217
      1273 TGGGGGAGAGAGCTGACCCAGTGAAGCAAGTGTCTTGAAGTCTTGTCTTACCTCTCG 1332
      218 GlyLysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
      1333 GGAGAACCAACCGGAATGGTCAAAAGTACTGGCTGCTGCTGGAGACAT---GAGCCC 1389
      231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe 250
      1390 CTGGCGGCTGAGACACGCTGCTGAAGGAGAGCTGCTGGGGGTGAACCGGACGAGCATC 1449
      251 LeuThrIleAsnSerGlnProAlaValAsnGluLysSerAspSerProThrValGly 270
      1450 CTCACCATTAACCTACAGGCCCAACATCAACAGGGAAGCCGCTCCGACCACTCTGGGC 1509
      271 TrpGlyGlyProGlyGlyTyrValTyrGluLysAlaTyrValGluPhePheCysSerLys 290
      1510 TGGGGCCCAAGCGGGGGCTATGTTCCAGAGGCGCTAAGATTTCACCTCCCGC 1569
      291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
      1570 GAGACAGCGGAGACACTCTCTGCAAGTGTGAAGAGTACGAGCTCCGGGTTAATTACAC 1629
      310 AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAlaVal 329
      1630 CTTCGCAATGTAAGGGGTAACAAATCAACCAATGCCCTGAACCTGACCCGAAATGCTGTC 1689
      330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
      1690 ACTTGGGCGATCTCTCCGTGGGAGATCAATCCAGCCACCGATGTGATCCGCTGAGC 1749
      350 PheAsnValIleTrpLysAspGluAlaPheGluIleTyrPheSerArgGlyTyrPalaSerLeuTyr 369
      1750 TTCATGTCTCTGGAGAGAGGAGGCTTGTGCTGATGAGCGGTGAGGAAACCTGTAT 1809
      370 ProGluAspGluAlaSerArgLysLeuValGluValGlyGlySerHisPheLeuVal 389
      1810 GAGGAGAGAGTCCCGCTCCGACCATCATTCACGTACATCCACGACATCTTCTCTGCTC 1869
      390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
      1870 AACCTGTGTGACAAATGACTTCCACTGACAAATGCTGTGCAAGTGTGAGAGAC 1926
```

RESULT 8
AAD20463

```
ID      AAD20463 standard; cDNA; 2219 BP.
XX
AC      AAD20463:
XX
DT      03-JAN-2002 (first entry)
XX
DE      Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #2.
XX
KW      Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
KW      antisense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
KW      pancreas; lung; brain; cytostatic; colon carcinoma; breast; colorectal;
KW      neuroblastoma; leukaemia; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      13..1983
FT      /*tag= a
FT      /product= "Human methylenetetrahydrofolate reductase
FT      (MTHFR) protein"
FT      /EC_number= "1.5.1.20"
XX
PN      US2001025030-A1.
XX
PD      27-SEP-2001.
XX
PF      01-DEC-2000; 2000US-0728910.
XX
PR      01-MAR-1999; 99US-0258928.
XX
PA      (ROZEN/) ROZEN R.
XX      (SEKH/) SEKHON J.
XX
PI      Rozen R, Sekhon J:
XX
DR      WPI: 2001-638509/73.
XX      P-PSDB: AAE12607.
XX
PT      New antisense nucleic acids, which are methylenetetrahydrofolate
PT      reductase inhibitors, useful for treating, stabilizing or preventing
PT      cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
PT      neuroblastoma
XX
PS      Disclosure; Fig 6; 68pp; English.
XX
CC      The invention relates to a non allele-specific antisense nucleic acids,
CC      which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
CC      nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
CC      catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
CC      5-methylenetetrahydrofolate, a co-substrate for methylation of
CC      homocysteine to methionine. The invention provides potential therapy for
CC      individuals with MTHFR deficiency. The non allele-specific antisense
CC      nucleic acids are useful for treating, stabilizing or preventing cancer,
CC      particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
CC      lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,
CC      neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
CC      therapy. The present sequence is human methylenetetrahydrofolate
CC      reductase (MTHFR) protein cDNA #2. Human MTHFR gene is mapped to
CC      chromosome 1p36.3.
XX
SQ      Sequence 2219 BP; 501 A; 657 C; 620 G; 441 T; 0 other:
XX
Alignment Scores:
Pred. No.: 5 19e-91 Length: 2219
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 22 Gaps: 7
US-09-720-451-6 (1-408) x AAD20463 (1-2219)
OY      1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
```

```

Db 676 GCCGATTTCATACACGACGCTTTCTTGAGGCTGACACACTTCTCCGCTTGTGAAG 735
QY 21 ASPCARGLInleGlylleThrcysProIleValProGlyIleMetProIleasn 40
Db 736 GCATGCACCGCATGGGACATCACTTCCCATGCTCCCGGGAGCTTCCATCCACAGGC 795
QY 41 TyrlcyslypheIleArghmetThrcylpheCysIlyThrIlyIleProIleAsnIleMet 60
Db 796 TACCACTCCCTTGCGCAGCTTGGAAGCTGTCCAGCTGGAGGTCGACAGACGAGCAAG 855
QY 61 AlaIleuGluProIleIleAspAsnGluGluIleValIleAspIleGlyIleHisIleu 80
Db 856 GACCTGATTGACCAATCAAGACAGATGCTGCCATCCGCACTATGATGACAGACTG 915
QY 81 GlyThrGluMetCysIlyIleIleuAlaHisGlyIleIleIysThrIleHisIleuThr 99
Db 916 GCGGTGAGCTGTGGCAGGAGCTTGTGGCCAGTGGCTTGGTCCAGGCTTCCACTTCTAC 975
QY 100 ThrIleuAsnMetGluIlySerAlaIleuAlaIleuMetAsnIleuGlyIleuIleGlu 119
Db 976 ACCCTCAACCGCAGATGGCTTACACAGAGGTCGAAGCGCTGGGATGTGAGACTGAG 1035
QY 120 SerIlyValSerIleSerIleuProIleParIleProIleAsnValIlePheArgValIleGlu 139
Db 1036 GACCC---AGGCTCCCTACCTGCTGCTCAGTCCGACCCACCCAGCCGAGAGGAA 1092
QY 140 AspValArgProIlePheThrPalaAsnArgProIlySerIlyIleSerArgThrIleGly 159
Db 1093 GATGTACGTCCTCCATCTTGGGCTCCAGCAACCAAGATTCACCTACCCAGAGAG 1152
QY 160 TrpAspGlnThrProIleHisGlyArgTrpGlyAspSerCysAsnProSerIlyGlyAlaIleu 179
Db 1153 TGGAGCAGATTCCCTAACGGCCGCTGGGCAATTCCTCTCCCTCCCTTGGGAGACTG 1212
QY 180 SerAspTrpGln-----PheMetArgProArgAlaArgAspIlyIleuValIleGlu 197
Db 1213 AAGGACTACTACCTCTTCTACCTGAAGACAGCAAGTCCCCAGAGAGAGACTGTAAGATG 1272
QY 198 TrpAlaValIleProIleuIlySerValIleuAspIleuArgPheArgIleuThrCysIleu 217
Db 1273 TGGGGGAGAGAGCTGACCACTGAGCAAGTGTTCGAGTTCCTTCTTACCTCTCG 1332
QY 218 GlyIlyIleuArgSerAsn-----ProIleSerGluIleuAspIly 230
Db 1333 GGAAGAACCAACCGAGATGCTCACAAGTACGCTTGCCTCCCTCGAAGCAT---GAGCC 1389
QY 231 LeuGlnProGluThrIlyIleIleAsnGluIleuGluIleuValIleAsnThrIlyGlyPhe 250
Db 1390 CTGGCGGCTGAGACACCTGCTGAAGAGAGAGCTCTCGGGGTGACCCGCGAGGCTATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValIleAsnGlyIleuIlySerAspSerProThrValGly 270
Db 1450 CTCACCACTCACTACAGCCCAACATCAACGGGAAGCCCTCCCTCCGACCATCGTGGGC 1509
QY 271 TrpGlyIleProGlyIleValIleuIlyIleuValIleuValIleuValIleuValIleuVal 290
Db 1510 TGGGGCCCGCAGCGGGGCTATGCTTCGCAAGAGCCCTACTTGAATTTTCTACCTCCGC 1569
QY 291 GluIlyIleuAspAlaIleuValIleuIlyCysIlyAsp---ArgThrSerIleuThrIlyMet 309
Db 1570 GAGACAGCGAGACACTTCTGCAAGTGTGAAGATGACGACTCCGGGTGAATTAATTCAC 1629
QY 310 AlaIleuAsnIlyAspGlySerTrpIlySerAsnValIleGlyIleuThrAspValIleuAlaIle 329
Db 1630 CTTGTCAATGTGAAGGCTGAAGAACATCAACAGTCCGCTGAAGCTCAGACCGAAGCTGTC 1689
QY 330 ThrTrpGlyValPheProAlaIlyGluIleIleGlnProThrIleValIleAspProValSer 349
Db 1690 ACTTGGGAGCATCTTCCCTGGGAGAGATCAATCCAGCCACCGTAGTGAATCCGCTAGC 1749
QY 350 PheAsnValIleTrpIlyAspGluIlePheGluIleIleTrpSerArgGlyIlePalaSerIleuThr 369

```

```

Db 1750 TTCATGTTCTGGAGAGACGAGSCCTTGGCCCTGTGATGAGCGGTGGGAAGCTGTAT 1809
QY 370 ProGluAspGluAlaSerArgIlyIleuValIleGluIleuValIleGlySerHisPheIleuVal 389
Db 1810 GAGGAGGAGTCCCGCTCCGACCATCATCTCAGTACATCAAGACAACTACTTCTGTGTC 1869
QY 390 SerIleuValIleAspAsnAspIlyr---IleAsnGlyAspIleuPheAlaIleuValIleuAsp 407
Db 1870 AACCTGTGGACATGACTTCCCATGACACTGACACTGCTCTGGCAGGTGTGGAAGAC 1926

RESULT 9
AAT09694
ID AAT09694 standard; cDNA; 2220 BP.
AC AAT09694;
XX
XX 15-OCT-1996 (first entry)
DE Human methylene-tetrahydrofolate-reductase cDNA.
KW Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;
KW cardiovascular disease; neurological disease; folic acid metabolism;
KW EC-1.5.1.20; enzyme; ss.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT mat_peptide 1..1980
FT /*tag= a

W0533054-A1.
PD 07-DEC-1995.
XX
XX 25-MAY-1995; 95KW-CA00314.
XX
XX 26-MAY-1994; 94GB-0010620.
XX
XX (UYWC-) UNIV MCGILL.
PA
PI Goyette P, Rozen R;
XX
XX WPI: 1996-030565/03.
DR P-PSDB: AAT09694.
XX
XX Human methylene-tetrahydrofolate reductase cDNA probe - for
PT detection of sequence abnormalities in methylene-tetrahydrofolate
PT reductase e.g. in cardiovascular, neurological or folic acid
PT metabolism disorders
XX
XX Disclosure: Fig. 1A-1F; 66pp; English.
XX
XX This sequence encoding human MTHFR has been localised to chromosome
CC 1p36.3. It may be used for the construction of DNA probes which may
CC be used for the identification of sequence abnormalities in patients
CC with severe or mild MTHFR deficiency. The resulting probe may also
CC be used in gene therapy to produce the MTHFR protein.
XX
SQ Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other;

Alignment Scores:
Pred. No.: 5.19e-91 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 17 Gaps: 7

US-09-720-451-6 (1-408) x AAT09694 (1-2220)
QY 1 AlaArgValIleValIleuValIleuPheThrAspThrAspIlePheIleuIlySerPheValAsn 20
Db 676 GCCGATTTCATACACGACGCTTTCTTGAGGCTGACACACTTCTCCGCTTGTGAAG 735

```

[illegible]

| | | | | |
|--|--|-----------------------|---|------|
| Oy | | 370 | ProglutinspiGlnAlaSerArgLysLeuValGluGluValGlySylSerHisPheLeuVal | 389 |
| Dd | | 1810 | GAGGAGAGACGCCCGCTCCCGCAGCATTCATCACTCACGCACGACACTACTTCTGGGT | 1869 |
| Oy | | 390 | SerLeuValAspAsnASPtyr---ileasnGlyaspLeupephalaValphealaasp | 407 |
| Dd | | 1870 | AACCTGGTGGAACATGACTTCACCACCTGACACACTGCCTCTGGCAGGGTGGAAAGC | 1926 |
| Db | | | | |
| XX | | | | |
| XX | AAA50633 | ID | AAA50633 standard; cDNA: 2220 BP. | |
| XX | | | | |
| AC | AAA50633: | | | |
| XX | | | | |
| DT | 19-DEC-2000 | (first entry) | | |
| XX | | | | |
| DE | | | Human methylentetrahydrofolate reductase cDNA. | |
| XX | | | | |
| KW | Methylentetrahydrofolate reductase; MTHFR; human; folic acid; | | | |
| KM | cardiovascular disorder; cancer; neuroblastoma; | | | |
| KM | colorectal carcinoma; osteoporosis; neural tube | | | |
| KW | neurological disorder; gene therapy; diagnosis; chromosome 1p36.3; | | | |
| KM | ss. | | | |
| XX | | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| FH | Key | Location/Qualifiers | | |
| FT | CDS | 13..1983 | | |
| FT | | /*tag= a | | |
| FT | | /EC_number="1.5.1.20" | | |
| FT | mutation | /replac(167,A) | | |
| FT | | /*tag= b | | |
| FT | mutation | /replac(482,A) | | |
| FT | | /*tag= c | | |
| FT | mutation | /replac(559,T) | | |
| FT | | /*tag= d | | |
| FT | mutation | /replac(677,T) | | |
| FT | | /*tag= e | | |
| FT | mutation | /replac(692,T) | | |
| FT | | /*tag= f | | |
| FT | mutation | /replac(764,T) | | |
| FT | | /*tag= g | | |
| FT | mutation | /replac(792+1,A) | | |
| FT | | /*tag= h | | |
| FT | mutation | /replac(985,T) | | |
| FT | | /*tag= i | | |
| FT | mutation | /replac(1015,T) | | |
| FT | | /*tag= j | | |
| FT | mutation | /replac(1018,T) | | |
| FT | | /*tag= k | | |
| FT | mutation | /replac(1298,C) | | |
| FT | | /*tag= l | | |
| FT | mutation | /replac(1317,C) | | |
| FT | | /*tag= m | | |
| XX | | | | |
| PN | WO200052205-A2. | | | |
| XX | | | | |
| PD | 08-SEP-2000. | | | |
| XX | | | | |
| PE | 28-FEB-2000; 2000WO-IB00442. | | | |
| XX | | | | |
| PR | 01-MAR-1999; 99US-0258928. | | | |
| XX | | | | |
| PA | (UYMC-) UNTV MCGILL. | | | |
| PI | | | | |
| Rozen R, | Goyette P; | | | |
| XX | | | | |
| DR | WP1: 2000-572192/53. | | | |
| DR | P-PDB; AA96186. | | | |
| XX | | | | |
| cDNA probe for the human methylentetrahydrofolate reductase (MTHFR), | | | | |
| useful in gene therapy and for diagnosing or treating MTHFR deficiency | | | | |
| PT which is associated with cardiovascular disorders or cancer - | | | | |

XX Claim 1; Fig 1A-F; 93pp; English.

XX The present sequence is that of cDNA coding for human
XX methylenetetrahydrofolate reductase (MTHFR, see AA196186), an enzyme
XX catalysing the NADPH-linked reduction of 5,10-methylenetetrahydrofolate
XX to 5-methyltetrahydrofolate, a co-substrate for methylation of
XX homocysteine to methionine. To identify human MTHFR cDNA, PCR
XX primers based on the porcine sequence were used to screen a human
XX liver lambda-gt10 cDNA library by PCR. A 1266 bp fragment was
XX obtained, and this was used to screen a human colon carcinoma cDNA
XX library to obtain the 2.2 kb clone. The MTHFR gene (see AA05654)
XX maps to chromosome 1p36.6. A cDNA probe for human MTHFR, which
XX hybridises to the present sequence, is claimed. This probe can be
XX used to identify MTHFR sequence abnormalities in individuals with
XX severe or mild MTHFR deficiency. These abnormalities may comprise
XX a mutation selected from 167G to A, 482G to A, 559C to T, 677C to
XX T, 692C to T, 764C to T, 792+1G to A, 985C to T, 1015C to T, 1081C
XX to T, 1298A to C and 1317T to C. MTHFR deficiency may be associated
XX with a cardiovascular disorder, cancer (especially neuroblastoma or
XX colorectal carcinoma), osteoporosis, neural tube defect in an
XX offspring of a patient, neurological disorders, and other disorders
XX influenced by folic acid metabolism. Also claimed are methods for
XX treating MTHFR deficiency by gene therapy or by administration of
XX MTHFR protein. Cancer can be treated by inhibiting MTHFR gene
XX expression or MTHFR protein activity, or by administering an agent
XX that modifies MTHFR gene expression.

SQ Sequence 2220 BP; 501 A; 658 C; 620 G; 441 T; 0 other;

Alignment Scores:

Pred. No.: 5,19e-91 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 21 Gaps: 7

US-09-720-451-6 (1-408) x AA050633 (1-2220)

QY 1 AAlaAGValIleValIleThGlnLeuPhetYrAspIlePhetLeuSphValAsn 20
DB 676 GCCGATTCATCATCACCACGCTTTCTTGAAGCTGACACATCTTCCGCTTGTGAAG 735
QY 21 AspCAsrAGlnIleGlyIleThCysProIleValProGlyIleMetProIleAsn 40
DB 736 GCATGCACCGACATGGCATCTTGCCTCCGCCGATCTTCCCATCCAGGCG 795
QY 41 TTTTysGlyPheIleArgMetThGlyPheCysIysThrIysIleProAlaSerIleMet 60
DB 796 TTACCACTCCCTTGGCAGACCTTGTGAAGCTGTCAAGCTGGAGGTCACAGAGATCAAG 855
QY 61 AAlaAlaLeuGluProIleLysAspAsnGluIuAlaValLysAlaTyrGlyIleHisLeu 80
DB 856 GACGATGATGAGCCAAATGAAGACACATGCTGCCATCCGCAATGTGGCATCGAGCTG 915
QY 81 GltYrHglMetCysLysLysIleLeuAlaHisGly--IleGlySphLeuHisLeuTyr 99
DB 916 GCGGTAGAGCTGTGCGCAGAGACTTGTGGCAGTGTGGTGGCAGAGGCTCCACTGTAC 975
QY 100 ThrIleAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnGluLysIleGluGlu 119
DB 976 ACCGTAAACCGGAGATGAGTACACAGAGAGTGTGAGAGCCGCTGGGAGATGTGAGTGA 1035
QY 120 SerLysValSerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu 139
DB 1036 GACCCC--AGGCTGCCCTTACGCTGGGCTCTCAGTGGCCACCCGACGCGGAGAGAA 1092
QY 140 AspValAlaArgProIlePheThrAlaAsnArgProLysSerTyrIleSerAlaGlyThrIleGly 159
DB 1093 GATGTAAGTCCCATCTTGTGGGCTCCGACACCAAGATTCATCTACCGGACCCAGAG 1152
QY 160 TrrAspGlnTyrProHisGlyAlaGlyTrpGlyLysSerCysAsnProSerTyrGlyAlaLeu 179

DB 1153 TGGACAGAGTTCCTTACAGCGCTGGGCAATTCCTCTCCCTCCCTTGGGAGAGTG 1212
QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysIysLeuValGluGlu 197
DB 1213 AAGCACTACTACTCTTCTTACCTGAAGACAACTCCGCCAAGGAGGATGGCTGAAGATG 1272
QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGlnArgPheArgIleuTyrCysLeu 217
DB 1273 TGGGGGGAGAGCTGACACAGAGAAAGCAAGTCTTGAAGTCTTGTCTTACCTCTCG 1332
QY 218 GltYsLysLeuArgSerAsp-----ProArgSerIleuLeuAspGly 230
DB 1333 GGAACACCAACCGGAAAGTGTCAAAAGTACTGCTGCTGCGAGACCAT--GAGCCC 1389
QY 221 LeuGlnProGluThrIysIleIleAsnGluGlnLeuGluLysIleAsnThrIysGlyPhe 250
DB 1390 CTGGGGGCTGAGACACACCTGCTGAAGAGAGAGCTGTGGGGTGAACCGCAGGCGATC 1449
QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyGlyLysSerAspSerProThrValGly 270
DB 1450 CTCACCATCAACTCACAGCCCAACATCAAGGGAAGCGCTCCGACCCCATCTGCGGC 1509
QY 271 TrpGlyIleProGlyIleTyrValTyrGlnLysAlaTyrValGluPhePheCysSerIys 290
DB 1510 TGGGGCCCGACGCGGGGCTATGCTTCCAGAGGCGCTACTTAAGTTTTCACTCCCGC 1569
QY 291 GltYsIleuAspAlaLeuValAspLysCysLysAsp---ArgThrSerIleuThrTyrMet 309
DB 1570 GAGACAGCGGAGAGCACTTGTCAAGTGTCAAGATGACAGATGAGAGCTCGGGTTAATTCAC 1629
QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
DB 1630 CTGTGCATGTGCAAGGAGTGAAGAAACATCAACATCCCTGATGACGCGGATGCTGTG 1689
QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
DB 1690 ACTTGGGCACTCTCCCTGGGCGAGAGATCATCCAGGCCACCGTAAGTGTCCCGTCAAG 1749
QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTrpSerArgGlyTrpAlaSerLeuTyr 369
DB 1750 TTCAATGCTCTGGAAGAGAGAGAGGCTTTGCTGTGATGACGGGTGGGAAGCTGTAT 1809
QY 370 ProGluAspGluAlaSerArgLysLeuValGluIuValGlyLysSerHisPheLeuVal 389
DB 1810 GAGGAGAGTCCCGCTCCGACCATCATTCACATGACACACACAACTACTTCTGTGTC 1869
QY 390 SerIleuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
DB 1870 AACCTGTGTGAACAATGACTTCCCACTGTGACAACTGCTGTGGAGGTGGAGAGAC 1926
RESULT 11
AAD20462
ID AAD20462 standard; cDNA; 2220 BP.
XX
AC AAD20462;
XX
DT 03-JAN-2002 (first entry)
XX
DE Human methylenetetrahydrofolate reductase (MTHFR) protein cDNA #1.
XX
KW Human; methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
XX antisenese therapy; BC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
KW pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;
XX neuroblastoma; leukaemia; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 1..1983
FT /tag= a
FT /product= "Human methylenetetrahydrofolate reductase
FT (MTHFR) protein"

FT /note= "CDS does not include start codon"
 FT /EC-number= "1.5.1.20"
 FT /partial
 XX
 XX
 XX US2001025030-A1.
 XX
 XX 27-SEP-2001.
 XX
 XX 01-DEC-2000; 2000US-0728910.
 XX
 XX 01-MAR-1999; 99US-0258928.
 XX
 XX (ROZEN/) ROZEN R.
 XX (SEKH/) SEKHON J.
 XX
 XX Rozen R, Sekhon J;
 XX
 XX MPI: 2001-638509/73.
 XX P-PSDB: AAE12606.
 XX
 XX New antisense nucleic acids, which are methylenetetrahydrofolate
 PT reductase inhibitors, useful for treating, stabilizing or preventing
 PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
 PT neuroblastoma -
 XX
 XX Disclosure: Fig 1; 68pp; English.
 PS
 XX The invention relates to a non allele-specific antisense nucleic acids,
 CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
 CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
 CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
 CC 5-methylenetetrahydrofolate, a co-substrate for methylation of
 CC homocysteine to methionine. The invention provides potential therapy for
 CC individuals with MTHFR deficiency. The non allele-specific antisense
 CC nucleic acids are useful for treating, stabilizing or preventing cancer,
 CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
 CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer,
 CC neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
 CC therapy. The present sequence is human methylenetetrahydrofolate
 CC reductase (MTHFR) protein cDNA #1. Human MTHFR gene is mapped to
 CC chromosome 1p36.3.
 XX
 XX Sequence 2220 BP; 501 A; 658 C; 620 G; 441 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 5,19e-91 Length: 2220
 Score: 979.00 Matches: 188
 Percent Similarity: 63.728 Conservative: 79
 Best Local Similarity: 44.878 Mismatches: 138
 Query Match: 44.764 Indels: 14
 DB: 22 Gaps: 7
 US-09-720-451-6 (1-408) x AAD20462 (1-2220)
 QY 1 AAlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
 DB 676 GCCGATTCATCATCAGCAGCTTCTTCTTGAGGCTGACATCTCTCCGCTTGGAG 735
 QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProIleIleMetProIleAsn 40
 DB 736 GCATGCAACGACATGGGCACTGCTGCCATGCTCCCGGATCTTTCCATCCAGGCG 795
 QY 41 TyrIleGlyPheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMet 60
 DB 796 TACCACCTCCCTGGCAGCTTGTGAAGCTTCCAAGCTGAGGTGGCAGGAGATCAAG 855
 QY 61 AAlaIleuGlnProIleLysAspAsnGluGluAlaValLysAlaTyrGlyIleHisLeu 80
 DB 856 GACGTGATGTAGCAATCAAGCAACAGATGTCGCATCCGCAACTATGGCATCGAGCTG 915
 QY 81 GlyThrGluMetCysLysLysIleLeuAlaHisGly---IleLysThrIleHisLeuTyr 99
 DB 916 GCCGTGAGCCCTGTGCCAGGAGCTTGTGCCAGTGGCTTGGTCCAGGCGCTCCATTCTAC 975

QY 100 ThrIleuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyIleGlu 119
 DB 976 ACCCTCAACCGCAGAGATGCTACACAGACAGATCTGAAGCGCTGGAGATGGACTGAG 1035
 QY 120 SerLysValSerArgSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu 139
 DB 1036 GACCC---AGCGCTGCCCTACCCCTGGCGTCTCAGTGGCCACCCCAAGCCGAGAGAA 1092
 QY 140 AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly 159
 DB 1093 GATGTAGCTGCCATCTTCTGGGCTCCAGACCAAGAGTTACATCTACCTACCCAGGAG 1152
 QY 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
 DB 1153 TGGGAGAGTTCCTTACAGGCGCGTGGGCAATTCCTTCCCTGCTTGGGAGGCTG 1212
 QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGlu 197
 DB 1213 AAGGACTACTACCTCTTCTTACAGCAAGATCCCCCAAGGAGAGGCTGTAAGATG 1272
 QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLysLeu 217
 DB 1273 TGGGGGAGAGCTGACCAAGTGAACAAAGTGTCTTGAAGCTTTGTCTTACCTGTC 1332
 QY 218 GlyLysLeuArgSerAsn-----ProTyrSerGluLeuAspGly 230
 DB 1333 GAGAGACCAACCGGAAATGGTCACAAAGTACTGCTGCCCTGGAGACGAT---GAGCCC 1389
 QY 221 LeuGlnProGluThrLysIleIleAsnGluGlnLeuLysIleAsnThrLysGlyPhe 250
 DB 1390 CTGGCGGCTGAGCAGCGCTGCTGAAGAGAGCTGCGGGGAAACCCGAGGGCATC 1449
 QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270
 DB 1450 CTCACATCACTACACAGCCCAACATCAACGGAAGCGTCTCCGACCCCATCGTGGCG 1509
 QY 271 TrpGlyLysProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
 DB 1510 TGGGCGCCACCGGGGGCTATGTCTCCAGAAAGCCCTACTTGAAGTCTTTCACCTCCGC 1569
 QY 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerIleThrTyrMet 309
 DB 1570 GAGACAGCGGAACCTTCTGCAAGTGTGAAAGATGACAGCTCCGGGTATATTACAC 1629
 QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
 DB 1630 CTGTGCATGTGAAGGGGTGAAMAACATGACCAATGCCCTGAACTGACAGCCGATGTGC 1689
 QY 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
 DB 1690 ACTTGGGCATCTTCTCGGGGAGAGATCATGACAGCCACCGTATGATCCCGTACGC 1749
 QY 350 PheAsnValTrpLysAspGluAlaPheGluIleTyrSerArgGlyTrpPalaSerLeuTyr 369
 DB 1750 TTCATGTTCTGGAAGAGAGAGCGCTTTCCTGCTGATTTGAGCGGTGGGAAACCTGTAT 1809
 QY 370 ProGluAspGluAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal 389
 DB 1810 GAGGAGAGAGTCCCGCTCCGACCATCCAGTACATCCAGCAACTACTCTCTGTGTC 1869
 QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
 DB 1870 AACCTGTGGACAAAGTCTTCCACTGGACAACTGCTCTGTGAGTGTGGAAAGAC 1926
 RESULT 12
 ID ABK13501 standard; cDNA: 2220 BP.
 AC ABK13501;
 XX 09-APR-2002 (first entry)
 XX

| US-09-720-451-6 (1-408) x ABRK13501 (1-2220) | |
|---|------|
| Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuPheValAsn | 20 |
| Db 676 GCCGATTCATCATCAGCAGCGCTTTCTTGGAGGCTGACACATTCCTCCGCTTGTGAAG | 735 |
| Qy 21 AspArgArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsn | 40 |
| Db 736 GCATGACCCGACGAGGGCAGTACCTTGCCCATGTGCCGGAGTCTTCCATCCAGGCC | 795 |
| Qy 41 TyrIysGlyPheIleArgMetThrGlyPheCysIleThrIleProAlaAspIleMet | 60 |
| Db 796 TACACACCCCTGGGCGAGCTTGTAAGCTGTCCAACTGGAGGTGCAAGAGATTCAG | 855 |
| Qy 61 AlaAlaLeuGlnProIleIlysAspAsnGluGlnAlaValIleLysAlaTyrGlyIleHisIleu | 80 |
| Db 856 GAGCTGATTTGGCCAAATCAAAAGAACAGAGATGCGCATCGCACTATGCGATCGAGCTG | 915 |
| Qy 81 GlyThrGlnMetCysIleLysIleLeuAlaHisGly---IleIysThrIleHisIleuTyr | 99 |
| Db 916 GCCGTGAGCCTGTCCACAGAGCTTGTGGCAGTGGCTTGGTGGCAGGCTCCACTTTCAC | 975 |
| Qy 100 ThrLeuAsnMetGlnIlysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlnIu | 119 |
| Db 976 ACCCTCAACCGCAGATGGCTACACACAGAGTGTCTGAAGCGCTGGGATGTCGACTGAG | 1035 |
| Qy 120 SerIysValSerArgSerLeuProThrPArgArgProAlaAsnValPheArgValIysGlu | 139 |
| Db 1036 GACCCC---AGGCGTCCCTTACCTGGGCTCTCAGTGCACCCCAAGCGCGAGAGGAA | 1099 |
| Qy 140 AspValArgProIlePheThrPalaAsnArgProIysSerTyrIleSerArgThrIleGly | 159 |
| Db 1093 GATGAGCTCCCATCTTCTGGGCTCCAGACCAAGATGATACGTACCGTACCCAGAG | 1159 |
| Qy 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu | 179 |
| Db 1153 TGGGCGAGTTCCCTAACGGCGGTGGGGAATTCCTTCCCTCCCTTGGGGAGCTG | 1213 |
| Qy 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspIysIleuValGluGlu | 197 |
| Db 1213 AAGGACTACTACCTTCTTACTGTAAGAGCAAGTCCCCCAAGAGAGAGCTGCTGAAGATG | 1273 |
| Qy 198 TrpAlaValProLeuIysSerValGluAspIleTyrGlnArgPheArgLeuTyrCysIleu | 217 |
| Db 1273 TGGGGGAGGAGCTGACACAGTGAAGAGTCTTGAAGTCTTCTTACTTACCTCTCG | 1333 |
| Qy 218 GlyIysIleuArgSerAsn-----ProIriserGlnIleuAspGly | 230 |
| Db 1333 GGAGAACCAACGAGATGGTCAACAAAGTACTTGCCTGCCCTGGGAACAT---GAGCCC | 1383 |
| Qy 231 LeuGlnProGlnThrIlysIleIleAsnGluGlnLeuGlnIlysIleAsnThrIysGlyPhe | 250 |
| Db 1390 CTGGCGGCTGAGACAGCTCTCTGAAGAGAGAGTGTGGGGTGAACCGCCAGGCAATC | 1444 |
| Qy 251 LeuThrIleAsnSerGlnProAlaValAsnGlyGluIysSerIysPserProThrValGly | 270 |
| Db 1450 CTCACACATCAACTACACACCCAAACATACAGGGGAAGCGTCTCGACCCCATCTGGGGC | 1500 |
| Qy 271 TrpGlyIysProGlyIleTyrValTyrGlnIleLysAlaTyrAlaIlePhePheCysSerIys | 290 |
| Db 1510 TGGGGCCCCAGGGGGGTATCTTCCAGAAGGCCCTACTTAAGTTTTCACCTTCCGC | 1566 |
| Qy 291 GlnIysIleuAspAlaLeuValAspIysCysIysAsp---ArgThrSerLeuThrTyrMet | 309 |
| Db 1570 GAGACAGCGGACACTTCTTCCAGTCTGTAAGAGTACAGAGCTCCGGTAAATTCACAC | 1623 |
| Qy 310 AlaValAsnIysAspGlySerTyrPlySerAsnValGlyGlnThrAspValAsnAlaVal | 329 |
| Db 1630 CTTCGTCATGTGAAGGGTGAAGAAATCATCACCAATGCCCTTAACCTGCAGCGCAATGCTGC | 1681 |
| Qy 330 ThrTrpGlyValPheProAlaIleGluIleIleGlnProThrIleValAspProValSer | 349 |


```

Db 1273 TGGGGGAGGAGGAGCTGACCATGTAAGCAAGTGTCTTGAACATCTTTGTCTTACCTCTCG 1332
Qy 218 GLYLYSLenhrArgSerAsn-----ProTPrSerGluLeuAspGly 230
Db 1333 GGAGAACCAACCAACCGGATGTCACAAAGACTTGTCCCTGCCGTGAACGAT--GAGGCC 1389
Qy 231 LeuGlnProGluThrLysIleIleAsnGluGlnLeuGlnLysIleAsnThrLysGlyPhe 250
Db 1390 CTGGCGGCTGAGACCACTGCTGAAGAGAGAGCTGCTGGGTGAACCCCGAGGGGATC 1449
Qy 251 LeuThrIleAsnSerGlnProAlaValAsnGlyGluLysSerAspSerProThrValGly 270
Db 1450 CTCACCACTCAACTACAGCCCAACATCAAGGGAGAGCCGCTCCGACCCCATCGTGGCC 1509
Qy 271 TrpGlyGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPhePheCysSerLys 290
Db 1510 TGGGGGCCCCAGGGGCTATGTTCTTCCAGAGGCGCTACTTGAAGTCTTTTCACTTCCCGC 1569
Qy 291 GluLysLeuAspAlaLeuValAspLysCysLysAsp--ArgThrSerLeuThrTyrMet 309
Db 1570 GAGACAGCGGAGAGCAGCTTGTGCAAGTGTGAGAGTACGAGCTCCGGGTAAATTACAC 1629
Qy 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
Db 1630 CTTTCACTATGTAAGAGGTGAAAACATCAACCAATCCCTGACATGCGACGCCGATGCTGC 1689
Qy 330 ThrTrpGlyValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer 349
Db 1690 ACTTGGGCGATCTTCCCTGGGCGAGATCATCCAGCCGCCAGTGGATGCCCGTGCAGC 1749
Qy 350 PheAsnValTrpLysAspGluAlaPheGluIleThrSerArgLysTrpAlaSerLeuTyr 369
Db 1750 TTTCATGTTCTGGAAGAGCAGAGCCCTTTGCCCTGGGATTTGAGCGGTGGGAAAGCTGTAT 1809
Qy 370 ProGluAspGluAlaSerArgLysLeuValGluGlnValGlyLysSerHisPheLeuVal 389
Db 1810 GAGAGAGAGTCCCGTCCCGACCATCATCCAGTACATCCAGCACGACCACTACTTCTGTC 1869
Qy 390 SerLeuValAspAsnAspTyr--IleAsnGlyAspLeuPheAlaValIleAlaAsp 407
Db 1870 AACCTGTGTGACATGACTTCCATCGACCAACTGCTCTGGCAGGTGTGGAGAGAC 1926

RESULT 14
ABK13540
ID ABK13540 standard; cDNA; 2220 BP.
XX
AC ABK13540;
XX
DT 09-APR-2002 (first entry)
XX
DE Human methylenetetrahydrofolate reductase (MTHFR) GA82A allele.
XX
KW Methylenetetrahydrofolate reductase; MTHFR; neuroleptic;
KW neuroprotective; tranquilizer; nootropic; antidepressant;
KW anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis;
KW anxiety; dementia; depression; epilepsy; Huntington's disease;
KW migraine; demyelinating disease; multiple sclerosis; pain;
KW Parkinson's disease; psychosis; stroke; cardiovascular disorder;
KW cancer; osteoporosis; metabolic disease; endocrine disease;
KW inborn error of metabolism; inflammation; immune disorder; human;
KW neoplasia; colorectal carcinoma; neoplastic disease; renal disease;
KW ss; EC number 1.5.1.20; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH CDS 1..1983
FT /*tag= a
FT /product= "MTHFR"
FT /note= "Methylenetetrahydrofolate reductase"
FT mutation replace(482,5)

```

```

FT XX /*tag= b
PN XX WO200196598-A2.
XX
PD 20-DEC-2001.
XX
PF 12-JUN-2001; 2001WO-CA00867.
XX
PR 12-JUN-2000; 2000US-0592595.
XX
PA (UYMC-) UNTV MCGILL.
XX
PI Rozen R;
XX
DR WPT. 2002-130741/17.
XX
PT P-PSDB; AA075422.
XX
PT Diagnosing subjects at risk for or suffering from a psychosis,
PT particularly schizophrenia comprises determining the presence of a
PT heterozygous methylenetetrahydrofolate reductase mutant allele in the
PT subject -
XX
PS Claim 4; Page -; 102pp; English.
XX
CC The invention describes a method of diagnosing a psychosis in a subject,
CC or a risk for or propensity to psychosis in a subject comprising
CC determining the presence of a heterozygous methylenetetrahydrofolate
CC reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a
CC subject. The method is useful for diagnosing subjects at risk of, or
CC suffering from a psychosis, particularly schizophrenia but includes
CC manic-depressive disease, organic psychotic disorders, psychosis in
CC alcohol or drug intoxication, postinfection psychosis, postpartum
CC psychosis, senile psychosis, traumatic psychosis and acute idiopathic
CC psychotic illnesses. A new pharmaceutical composition is used for
CC treating amyotrophic lateral sclerosis, anxiety, dementia, depression,
CC epilepsy, Huntington's disease, migraine, demyelinating disease, multiple
CC sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or
CC stroke. Deficiency may be associated with diseases including
CC cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal
CC carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors
CC of metabolism, inflammation, immune disorders, neoplastic disease and
CC renal disease. This sequence encodes a mutant human
CC methylenetetrahydrofolate reductase (MTHFR, EC number 1.5.1.20), in which
CC nucleotide 482 has been altered to produce a mutant allele, described in
CC the method of the invention.
CC Note: This sequence does not appear in the specification but has been
CC created from the wild type sequence (ABK13501) using information given
CC in claim 4 of the invention.
XX
SQ Sequence 2220 BP; 502 A; 658 C; 619 G; 441 T; 0 other;

Alignment Scores:
Pred. No.: 5,19e-91 Length: 2220
Score: 979.00 Matches: 188
Percent Similarity: 63.72% Conservative: 79
Best Local Similarity: 44.87% Mismatches: 138
Query Match: 44.76% Indels: 14
DB: 24 Gaps: 7

US-09-720-451-6 (1-408) x ABK13540 (1-2220)
Qy 1 AlaArgValIleValThrGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
Db 676 GCCGATTTTCATATACAGCAGCTTTCTTGAAGCGACACATCTCTCCGTTTGAAG 735
Qy 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyTylIleMetProIleAsnAsn 40
Db 736 GCATGACCGACATGGGACATCTTCCCATCGTCCCGGATCTTCCCATCGAGGCC 795
Qy 41 TyrLysGlyPheIleArgMetThrGlyPheCysLysThrLysIleProAlaSpIleMet 60
Db 796 TACCACTCCCTTGGCAGCTGTGGAAGCTGTCCAGCTGAGAGTGCACAGAGATCAAG 855

```

| | | | |
|----|------|---|-------|
| QY | 61 | AlaAlaIleuInpProIleuLysAspAsnGluGluAlaValLysAlaIleIleHisLeu | 80 |
| | | | |
| Db | 856 | GACGTGATTGGACCAATCAAGAACAGATGCTGCATCCGGAACACTATGGATCCAGCTG | 915 |
| QY | 81 | glyThrIuMetCysLysLysIleLeuAlaHisGly---IleLysThrLeuHisLeuTyr | 99 |
| | | | |
| Db | 916 | GCCCGGACCCCTGTGCCAGAGAGCTTCTGGCCAGTGGCTTGTTGGTGCAGAGCCTCCACTTCTAC | 975 |
| QY | 100 | ThrLeuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGlu | 119 |
| | | | |
| Db | 976 | ACCCTCAACCCCGAGATGGGTACACACAGAGTGTCTAAACGGCTGGCGGATGGAGCTAG | 10359 |
| QY | 120 | SerLysValSerAspSerLeuProTyrArgArgProAlaAsnValPheArgValLysGlu | 139 |
| | | | |
| Db | 1036 | GACCC---AGCGTCCCCCTACCTGGGGCTCTAGTGGCCACCCCAAGCCCGGAGAGAA | 10922 |
| QY | 140 | AspValArgProIlePheThrPalaAsnArgProLysSerTyrIleSerArgThrIleGly | 159 |
| | | | |
| Db | 1093 | GATGACGTCCCACTCTCTGGGGCTCCAGACCAAGATTAATCATCTACCGTACCCAGAG | 11522 |
| QY | 160 | TrpAspGlnTyrProHisGlyArgTyrPrgLysAspSerCysAsnProSerTyrGlyAlaLeu | 179 |
| | | | |
| Db | 1153 | TGGGACGAGTTCCTTAACGGCCCTGGGGAAATTCCTTCCCTCCCTTGGGAGCTG | 1212 |
| QY | 180 | SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGlu | 197 |
| | | | |
| Db | 1213 | AAGGACTACTACCTCTCTACGTAAGAGCAAGTCCCAAGAGAGAGCTGTGAAGATG | 1272 |
| QY | 198 | TrpAlaValProLeuLysSerValGluAspIleTyrGluArgPheArgLeuTyrCysLeu | 217 |
| | | | |
| Db | 1273 | TGGGGGAGGAGCTGACGACATGAAGCAAGTCTTGAAGTCTTGTTCTTACCTCTCG | 13322 |
| QY | 218 | GlyLysLeuArgSerAsn-----ProTyrSerGluMetAspLys | 230 |
| | | | |
| Db | 1333 | GGAAGACCAACCGGATGTGTCAAAAGTACTGTGCTCTCCCTGGAAAGAT---GAGCC | 13899 |
| QY | 231 | LeuGlnProGluThrLysIleIleAsnGluGlnLeuGluLysIleAsnThrLysGlyPhe | 250 |
| | | | |
| Db | 1390 | CTGGCGGCTGAGACACCTGCTGGAAGAGAGACTCTCTGGGTGAACCGCAGGGCATC | 1449 |
| QY | 251 | LeuThrIleAsnSerGlnProAlaValAlaAsnGlyLysSerAspSerProThrValGly | 270 |
| | | | |
| Db | 1450 | CTCAACATCACTACAGCCCAATCAACAGGAAAGCCCTCTCCGACCCCATCGTGGC | 15099 |
| QY | 271 | TrpIleGlyProGlyGlyTyrValTyrGlnLysAlaTyrValGluPheCysSerLys | 290 |
| | | | |
| Db | 1510 | TGGGGCCCAAGCGGGGCTATGTCTTCCAGAAAGGCTTACTTGAATTTTCACTTCCGC | 1569 |
| QY | 291 | GluLysLeuAspAlaLeuValAspLysCysLysAsp--ArgThrSerLeuThrTyrMet | 309 |
| | | | |
| Db | 1570 | GAGCAACCGGAGACACTTGCACAGTGCAGAAGATAGACACTCCGGGTTAAATTACAC | 1629 |
| QY | 310 | AlaValAsnLysAspGlySerTyrLysSerAsnValGlyGlnThrAspValAlaValAla | 329 |
| | | | |
| Db | 1630 | CTTCTCAATGTGAAGGTTGAAGAAACATCAACCAATGCCCTGAACTGCACCCAAATCTGTG | 1669 |
| QY | 330 | ThrTrpIleValPheProAlaLysGluIleIleGlnProThrIleValAspProValSer | 349 |
| | | | |
| Db | 1690 | ACTGGGGGCACTTCTCCCTGGGCGAGAGATCAATCCAGCCACCGTATGTGATCCGCTAGC | 1749 |
| QY | 350 | PheAsnValTyrLysAspGluAlaPheGluIleTyrPserArgGlyTyrPalaSerLeuTyr | 369 |
| | | | |
| Db | 1750 | TTTCAATGTTCTGGAAAGCAAGGCGCTTGGCCCTGTGGATTTGACCGGTGGGAAACCTGAT | 1809 |
| QY | 370 | ProGluAspLysAlaSerArgLysLeuValGluGluValGlyLysSerHisPheLeuVal | 389 |
| | | | |
| Db | 1810 | GAGGAGAGTCCCGTCCGCGACATCAATCAATCAAGTCAATCAAGCAACTACTTCTGTGTC | 1869 |
| QY | 390 | SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp | 407 |
| | | | |
| Db | 1870 | AACCTGGTGGACATGACTTCCACATGGCAACCTGCTTGGCAGGTGTGGGAAGAC | 1926 |

| | |
|-----|---|
| ID | ABK13541 |
| AC | ABK13541 standard; cDNA: 2220 bp. |
| XX | |
| XX | ABK13541; |
| DT | 09-APR-2002 (first entry) |
| XX | |
| DE | Human methyltetrahydrofolate reductase (MTHFR) C559T allele. |
| XX | |
| KW | Methyltetrahydrofolate reductase; MTHFR; neuroleptic; |
| KW | neuroprotective; tranquiliser; nootropic; antidepressant; |
| KW | anticonvulsant; cerebroprotective; amyotrophic lateral sclerosis; |
| KW | anxiety; dementia; depression; epilepsy; Huntington's disease; |
| KW | migraine; demyelinating disease; multiple sclerosis; pain; |
| KW | Parkinson's disease; psychosis; stroke; cardiovascular disorder; |
| KW | cancer; osteoporosis; metabolic disease; endocrine disease; |
| KW | inborn error of metabolism; inflammation; immune disorder; human; |
| KW | neuroblastoma; colorectal carcinoma; neoplastic disease; renal disease; |
| ss | EC number 1.5.1.20; mutant. |
| XX | |
| OS | Homo sapiens. |
| XX | Synthetic. |
| FH | |
| FH | Key Location/Qualifiers |
| CDS | 1..1983 |
| FT | /tag= a |
| FT | /product= "MTHFR" |
| FT | /note= "Methyltetrahydrofolate reductase" |
| FT | replac(559,C) |
| FT | /tag= b |
| PN | |
| XX | WO200196598-A2. |
| PD | |
| XX | 20-DEC-2001. |
| PF | |
| XX | 12-JUN-2001; 2001WO-Ca00867. |
| PR | |
| XX | 12-JUN-2000; 2000US-0592595. |
| PA | (UYMC-) UNIV MCGILL. |
| XX | |
| PI | Rozen R; |
| DR | WPI: 2002-130741/17. |
| XX | P-PSDB: AAU75423. |
| PT | Diagnosing subjects at risk for or suffering from a psychosis, |
| PT | particularly schizophrenia comprises determining the presence of a |
| PT | heterozygous methyltetrahydrofolate reductase mutant allele in the |
| XX | subject - |
| PS | |
| XX | Claim 4; Page -: 102pp; English. |

The invention describes a method of diagnosing a psychosis in a subject, or a risk for or propensity to psychosis in a subject comprising determining the presence of a heterozygous methyltetrahydrofolate reductase (MTHFR) mutant allele in a nucleic acid sample obtained from a subject. The method is useful for diagnosing subjects at risk of, or suffering from a psychosis, particularly schizophrenia but includes manic-depressive disease, organic psychotic disorders, psychosis in alcohol or drug intoxication, postinfection psychosis, postpartum psychosis, senile psychosis, traumatic psychosis and acute idiopathic psychotic illnesses. A new pharmaceutical composition is used for treating amyotrophic lateral sclerosis, anxiety, dementia, depression, epilepsy, Huntington's disease, migraine, demyelinating disease, multiple sclerosis, pain, Parkinson's disease, schizophrenia, psychoses, or stroke. Deficiency may be associated with diseases including cardiovascular disorders, cancer (e.g. neuroblastoma and colorectal carcinoma), osteoporosis, metabolic or endocrine disease, inborn errors of metabolism, inflammation, immune disorders, neoplastic disease and renal disease. This sequence encodes a mutant human

nucleotide 559 has been altered to produce a mutant allele, described in

CC the method of the invention.
 CC Note: This sequence does not appear in the specification but has been
 CC created from the wild type sequence (ABK13501) using information given
 CC in claim 4 of the invention.

XX Sequence 2220 BP; 501 A; 657 C; 620 G; 442 T; 0 other;

Alignment Scores:

| Pred. No.: | 519e-91 | Length: | 2220 |
|------------------------|---------|---------------|------|
| Score: | 979.00 | Matches: | 188 |
| Percent Similarity: | 63.72% | Conservative: | 79 |
| Best Local Similarity: | 44.87% | Mismatches: | 138 |
| Query Match: | 44.76% | Indels: | 14 |
| | | Gaps: | 7 |

US-09-720-451-6 (1-408) x ABK13541 (1-2220)

```

QY 1 AlaATGValIleValThGlnLeuPheTyrAspThrAspIlePheLeuLysPheValAsn 20
   |||
   |||
   |||
Db 676 GCCGATTTCATCATCATCACGACCTTTCTTGAGGCTGCACACATCTCTCCGCTTGTGAG 735

QY 21 AspCysArgGlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAsn 40
   |||
   |||
   |||
Db 736 GCATCCACCGACACATGGGATCATCTTCCCATCGTCCCGGATCTTCCCATCCAGGCG 795

QY 41 TyrIlysglyPheIleArgMetThrGlyPheCysIlyThrIlyIleProAlaAspIleMet 60
   |||
   |||
   |||
Db 796 TACCACTCCCTTCGACACCTTGAGAGCTGTCCAAAGCTGAGGAGGCCACAGAAATCAAG 855

QY 61 AlaIleLeuGlnProIleLysAspAsnGlnGlnAlaValLysAlaIlyGlyIleHisLeu 80
   |||
   |||
   |||
Db 856 GACGGATTGAGCCATTAAGACACAGCATGCTCCATCCCAACTATGAGATCGAGCTG 915

QY 81 GlyThrIleMetCysLysIleLeuAlaHisGly--IleLysThrIleuHisLeuTyr 99
   |||
   |||
   |||
Db 916 GCCGTGACGCTGTGCAGACCTTCTGGCCAGTGGCTTGGCCAGGCTCCACTTCTAC 975

QY 100 ThrIleuAsnMetGluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyIleGluGln 119
   |||
   |||
   |||
Db 976 ACCCTCAACCCGAGATGGCTACCAAGAGGTGCTGAAGCGCTGGGAGATGTGAGCTGAG 1035

QY 120 SerLysValSerArgSerLeuProIleArgArgProAlaAsnValPheArgValLysGln 139
   |||
   |||
   |||
Db 1036 GACCC---AGCGTCCCTTACCTGAGGCTCTCAGTGCACCCCAAGCGCCGAGAGAA 1092

QY 140 AspValArgProIlePheThrAlaAsnArgProLysSerTyrIleSerArgThrIleGly 159
   |||
   |||
   |||
Db 1093 GATGTAGTCCCATCTCTGTGGCTCCAGAACAAAGATTACATCTACCTACCCAGAGAG 1152

QY 160 TrpAspGlnTyrProHisGlyArgTyrGlyAspSerCysAsnProSerTyrGlyAlaLeu 179
   |||
   |||
   |||
Db 1153 TGGAGACGAGTCCCTAACGGCGCTGGGCAATCTCTCCCTGCTTGGGAGAGCTG 1212

QY 180 SerAspTyrGln-----PheMetArgProArgAlaArgAspLysLysLeuValGluGln 197
   |||
   |||
   |||
Db 1213 AAGGACTACTACCTCTCTACTGTAAGACAAAGTCCCAAGAGGAGCTGTGAAGATG 1272

QY 198 TrpAlaValProLeuLysSerValGluAspIleTyrGlnArgPheArgLeuTyrCysLeu 217
   |||
   |||
   |||
Db 1273 TGGGGGAGAGAGCTGACCAAGCAAGGTCTTTGAAGTCTTTGTCTTACCTCTCG 1332

QY 218 GlyLysLeuArgSerAsn-----ProTrpSerGlnLeuAspGly 230
   |||
   |||
   |||
Db 1333 GGAACAACCAACCGAATGGTCAACAAGTACTGCTGCTGCAACGAT--GAGCC 1389

QY 231 LeuGlnProGlnTyrIleIleAsnGlnGlnLeuLysIleAsnThrLysGlyPhe 250
   |||
   |||
   |||
Db 1390 CTGGCGGCTGAGACACCGCTGTAAGAGAGAGCTGTGGGTGAACCGCAGGCGATC 1449

QY 251 LeuThrIleAsnSerGlnProAlaValAsnGlyLysSerAspSerProThrValGly 270
   |||
   |||
   |||
Db 1450 CTCACATCAACTACAGCCCAACATCAACGGGAGCGCTCTCCGACCCCATCGTGGCG 1509
  
```

```

QY 271 TrpGlyLysProGlyIlyTyrValTyrGlnLysAlaTyrValGlnPhePheCysSerLys 290
   |||
   |||
   |||
Db 1510 TGGGGCCCCAGCGGGGCTATGCTCTCCAGAAAGCCCTACTTACAGTTTCCACTCCCG 1569

QY 291 GlnLysLeuAspAlaLeuValAspLysCysLysAsp---ArgThrSerLeuThrTyrMet 309
   |||
   |||
   |||
Db 1570 GAGACAGCGGAAGGACTTCTGCAAGTGTCTGAAGAAGTACGAGCTCCGGTTAATTACAC 1629

QY 310 AlaValAsnLysAspGlySerTrpLysSerAsnValGlyGlnThrAspValAsnAlaVal 329
   |||
   |||
   |||
Db 1630 CTGTCAATGTGAAGGTGAAGAAACATCCACAAATGCCCTGAACTGCAAGCGAATGCTGTC 1689

QY 330 ThrTrpGlyValPheProAlaLysGlnIleIleGlnProThrIleValAspProValSer 349
   |||
   |||
   |||
Db 1690 ACTTGGGCACTCTTCCCTGGGCGAGATCATCAGCCACCGTAGTGTGATCCCGTCAGC 1749

QY 350 PheAsnValTrpLysAspGlnAlaPheGlnIleTrpSerArgGlyTrpAlaSerLeuTyr 369
   |||
   |||
   |||
Db 1750 TTCATGTCTGGAAGACAGAGGCTTTGCCCTGTGATGAGCGGTGGGAAAGCTGTAT 1809

QY 370 ProGlnAspGlnAlaSerArgLysLeuValGlnGlnValGlyLysSerHisPheLeuVal 389
   |||
   |||
   |||
Db 1810 GAGGAGAGAGTCCCGGTCCGACCATCATCATCATCATCATCATCATCATCATCATCAT 1869

QY 390 SerLeuValAspAsnAspTyr---IleAsnGlyAspLeuPheAlaValPheAlaAsp 407
   |||
   |||
   |||
Db 1870 AACCTGTGACATATACATCTCCACTGGAACAACTGCTGTGGAGGTGGTGAAGAC 1926
  
```

Search completed: February 14, 2003, 23:16:00
 Job time : 301 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 21:20:31 ; Search time 41 Seconds

(without alignments)
956.655 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVIYQLFYDIDIFLKEFN.....VSLVDNDYINGDLFAVADF 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1834 | 83.9 | 606 | 2 | T00696 |
| 2 | 1777 | 81.3 | 592 | 2 | T47821 |
| 3 | 895 | 40.9 | 603 | 2 | T38920 |
| 4 | 778 | 35.6 | 599 | 2 | S64136 |
| 5 | 711 | 32.5 | 615 | 2 | T15423 |
| 6 | 629 | 28.8 | 641 | 2 | T38659 |
| 7 | 578.5 | 26.5 | 657 | 2 | S63459 |
| 8 | 486 | 22.2 | 416 | 2 | S46454 |
| 9 | 421.5 | 19.3 | 566 | 2 | T42227 |
| 10 | 222 | 10.2 | 290 | 2 | F83591 |
| 11 | 222 | 10.2 | 296 | 2 | D70424 |
| 12 | 221 | 10.1 | 314 | 2 | C87514 |
| 13 | 215 | 9.8 | 275 | 2 | F82720 |
| 14 | 193 | 8.8 | 307 | 2 | T34973 |
| 15 | 189 | 8.6 | 292 | 1 | H64123 |
| 16 | 189 | 8.6 | 296 | 1 | RDECMH |
| 17 | 189 | 8.6 | 296 | 2 | E91237 |
| 18 | 189 | 8.6 | 296 | 2 | B86085 |
| 19 | 189 | 8.6 | 296 | 2 | AB0937 |
| 20 | 188 | 8.6 | 282 | 2 | D81326 |
| 21 | 184 | 8.4 | 296 | 1 | S03169 |
| 22 | 182 | 8.3 | 292 | 2 | F81880 |
| 23 | 182 | 8.3 | 292 | 2 | D81140 |
| 24 | 179 | 8.2 | 294 | 2 | AC0015 |
| 25 | 179 | 8.2 | 296 | 2 | AC2837 |
| 26 | 179 | 8.2 | 312 | 2 | E97614 |
| 27 | 171 | 7.8 | 304 | 2 | C82045 |
| 28 | 168 | 7.7 | 292 | 2 | A84935 |
| 29 | 163 | 7.5 | 283 | 2 | F86781 |

ALIGNMENTS

RESULT 1

T00696

Probable methylenetetrahydrofolate reductase [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F6E13.29

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00696; B84875

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke submitted to the EMBL Data Library, June 1998

A:Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.

A:Reference number: Z14180

A:Accession: T00696

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-606 <RDB>

A:Cross-references: EMBL:AC004005; NID:G3212846; PID:G3212869

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:1061197

A:Accession: B84875

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-606 <STO>

A:Cross-references: GB:AE002093; NID:G3212869; PIDN:AAC23420.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2G44160; F6E13.29

A:Map position: 155/1; 194/3; 252/3; 310/3; 344/2; 381/3; 408/3; 468/1; 579/3

A:Insertions:

Query Match 83.9%; Score 1834; DB 2; Length 606;
Best Local Similarity 82.1%; Pred. No. 1.3e-138;

Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 1 | ARVIYQLFYDIDIFLKEFNDCRQIGITCPYVGPINNYKGFIMTGFCKTKIPADIM | 60 |
| DB | 199 | ADIVITQLFYDIDIFLKEFNDCRQIGISCPYVGPINNYRQFLMTGFCFKIPVEVM | 258 |
| QY | 61 | AALPEIKDNEAVKAVGILHGTGEMCKKIIAHGKTHLVTKLNMWKSALAIIMNLGIIIES | 120 |
| DB | 259 | AALPEIKDNEAVKAVGILHGTGEMCKKIIAHGKTHLVTKLNMWKSALAIIMNLGIIIES | 318 |
| QY | 121 | KVSRSLEPWRPAPVNFVKEDVPRIFWANRPKSYISRTIGWDQYPHGRWDSQNPYSGLS | 180 |
| DB | 319 | KISRSLEPWRPAPVNFVKEDVPRIFWANRPKSYISRTIGWDQYPHGRWDSQNPYSGLS | 378 |
| QY | 181 | DYFMPRRARADKKLVEMAVPLKSVDIYERRLYCLGLRNPSELDGLQPETRIINE | 240 |
| DB | 379 | DHGFSPRRARADKKLVEMAVPLKSVDIYERRLYCLGLRNPSELDGLQPETRIINE | 438 |

OY 241 QLEKINTEGFLTINSQPVRNGEKSDSPVWGSGPGCGYYQKAAYVEFCSKEKLDALVFKC 300
|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||
Db 439 QLIVNNSGGLLTINSQPSVNAERSDSPVWGSGPGPYGYQAILEFFCSKERLDAVAKC 498

OY 301 KDRTSLIYMAVNKGDSMKSNVGQTIDVNAVTVGVPEPAKEIIPTIVDPVSFNWVKDEAFEI 360
|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||
Db 499 KALPSITVMANNKKGEQWNVNTAQADVNAVTVGWGFPAKEIIPDTIVDPASFNWKDEAFET 558

OY 361 WSRGNASTLYPDEDASRKLYEEVGSHEFLSYLVNDNYINGDLFAVPAD 407
|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||::|||||||::|||
Db 559 WSRWANLYPEADBSRNLLEEYKNSYLVSLEVNDYINGDLFAVPAD 605

RESULT 2
T47821
methyleletrahydrofolate reductase MTHFR1 - Arabidopsis thaliana
N:Alternate names: protein F2AG16.240
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_rev1sion 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47821
R:D'Angelio, M.; Yezi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z4477
A:Accession: T47821
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-592 <DAN>
A:Cross-references: EMBL:ALJ38647
C:Experimental source: cultivar Columbia; BAC clone F2AG16
A:Map position: 3
A:Introns: 155/1, 182/3, 240/3, 298/3, 332/2, 369/3, 396/3, 456/1, 567/3
A>Note: F2AG16.240

[illegible]

RESULT 3
138920
methylenetetrahydrofolate reductase 2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T38920
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z21817
A:Accession: T38920
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-603 <PEA>
A:Cross-references: EMBL:Z69728, PIDN:CA93581.1, GSPDB:GN00066; SPDB:SPAC56F8.10
C:Genetics:
A:Gene: SPDB:SPAC56F8.10
A:Map position: 1

```

Query Match          40.9% ; Score 895 ; DB 2; Length 603;
Best Local Similarity 45.2% ; Pred. No. 1,6e-63;
Matches 189; Conservative 57; Mismatches 148; Indels 24; Gaps 8;

QY      1 ARVITVLFYDTDLFLFKFVNDRCOIGTCPIVPGIMPINNYKGIKMTGCECKTKIPADIM 60
      1 | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      178 ADFIVTQGFYVDNFIWADVKRAAGINIPFPPIQIADSFIRAKKSGVYKIPQHF 237
      1 | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      61 AALEIKNEEKVAKYGIHILGTEMCCKTLANGCITHLHYLTNMEKSLALIMNLGLEES 120
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      238 DTLVAVKDDDESVRERGVSELLVEMCKRLIAGCITRLHETLYNLEKAKMIIERGLDEN 297
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      121 KYRSRLSPRRPANN-----FVKEDVPIPIPMANPKSYISRTIGMDQYPHGRMGDS 171
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      298 LA-----PIVDNNVBLTNASQDRINCGVPIFMRTRENSYVSRTQMDLPHGRMGDS 353
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      172 CNPSTGALSIOF---MRPRADKKLYEEMAVPLKSVEDIYERRLYCLGKLRNPNPSEL 228
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      354 RSPAGGEFDPIRYYGLRMSPP---KEITTSWSP--KSYSEIGDLEAFRCCKKISSLPMDL 408
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      229 DDLQETKIKINOLEKIKWTGFLTINSOPANNGEKSDPTVGMGCPGGYVQAKAYVEFC 288
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      409 P-ITSEADILIRQLLSMNRNAPFLTINSQPALNGEKSSHPVGPMPNGYVQKRYVEFV 467
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      289 SKEKLDALVDCRDKPTSLTYMAVNRKDGSMKSNVGOITVNAVTVGVFPKAKETIOTYIPV 348
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      468 HPSLLNLEKTEYFKKSNYSVSYFTVNRNGDLDITNSQETIAPNATWGVFPNRELIQPTIVEST 527
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      349 SPNWKMDFAFELTWSGMSGLYDEDEASRKYLYEEVGGSHFLSLVNDVPIIN-DLEAFV 405
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      528 SFLMADEAYTSL-GHEMANAYSPPDSISKRLVSMKKEVFLCYIVDNDPQNSQLFDVF 584
      1 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 4
 S64136
 methylentetrahydrofolate reductase (NADPH2) (EC 1.5.1.20) MET13 - yeast (*Saccharomyces Nalt*)
 N:Alternate names: protein G2882; protein YGL125w; ribosomal protein Yml45; mitochond
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 03-Jun-2002
 C:Accession: S64136; S53294; S78029; S72026
 R:Cerdan, E.; Rodriguez-Forres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64134
 A:Accession: S64136
 A:Molecule type: DNA
 A:Residues: 1-599 <CER>
 A:Cross-references: EMBL:D272647; NID:g1322686; PIDN:CA96633.1; PID:g1322687; MIPS:YG
 A:Experimental source: strain S288C
 R:Hosaka, K.; Nakawa, J.; Kodaki, T.; Ishizu, H.; Yamashita, S.
 J. Biochem. 116, 1317-1321, 1994
 A:Title: Cloning and sequence of the SCS3 gene which is required for inositol prototro
 A:Reference number: JX0365; MUID:95221330; PMID:7706223
 A:Accession: S53294
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-113 'KRRSS' <HOS>
 A:Cross-references: EMBL:D21200
 R:Itikawa, M.; Grack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wilt
 Eur. J. Biochem. 245, 449-456, 1997

A>Title: Identification and characterization of the genes for mitochondrial ribosomal p
A:Reference number: S78018; MUID:97296414; PMID:9151978
A:Accession: S78029
A:Molecule type: protein
A:Residues: 39-40; 'LA',43-47, 'A',49 <KIT>
A>Note: this protein was identified as ribosomal protein Yml45, mitochondrial
Ritzon, B.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Cadahia, J.L.; Cerdan, E.
Yeast 12, 1047-1051, 1996
A>Title: Identification of a putative methylenetetrahydrofolate reductase by sequence an
A:Reference number: S72026; MUID:97051592; PMID:8896269
A:Accession: S72026
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-72; 'R',74-599 <TIZ>
A:Cross-references: EMBL:X94106; NID:91628448; PIDN:CAA63833.1; PID:91628449
A>Note: this protein was identified as putative methylenetetrahydrofolate reductase
C:Genetics:
A:Gene: SGD:MET13; MET11; MRPL45
A:Cross-references: MIPS:YGL125W; SGD:S0003093
A:Map position: 7L
A:Genome: nuclear
C:Function: <METH>
A:Description: oxidoreductase
A>Note: this function seems to contradict the other function assigned to this protein
C:Function: <RIB>
A:Pathway: protein biosynthesis
A>Note: this function seems to contradict the other function assigned to this protein
C:Keywords: mitochondrion; oxidoreductase; protein biosynthesis; ribosome

Query Match 35.6%; Score 778; DB 2; Length 599;
Best Local Similarity 40.9%; Pred. No. 3.5e-54;
Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

4 IYVOLFVDFDIFLKVNDRCRIGITCPVPGIMPINNYKGFIMTGFCKTKIPADIM 63
DB 185 IITQMFYDNDLNLWCSOVRAGMDVPIPEIMPTTYAFLRRIQWGISIPQHRSS 244
QY 64 EPIKDNKNEAVKAYGIIHGTCEKCKILAHG-IKTLHLYTLMEKSALAILMNLGLI-EESK 121
DB 245 DPKDDDELIVRIGITNLIVEMCKRLDGSYVSHLYTMLEKAPMILIERILIPTESE 304
QY 122 VSRS-----LPMRRPANYRVEDVRPIFMANRPKSYSTRTGW--DQYPRGMRGDSQNP 175
DB 305 FNAHPLAVLPWRKSLNPKRKNEDVRPIFWRKRPSTYAKRSQVAFDFPNRGRGDSQNP 364
QY 176 YGAL-----SDYQFMRPRARDKLVEEAVPLKSVEDYERFRLYCLGKLSNFWSELDGL 231
DB 365 FGDLDCGSD--LIQSA--KCELEWSTP-TSINDVAFVINYLNGLNKLCPWSDIP-I 418
QY 232 OPEPTKILNQLEKINTKGFILINSOPAVNGEKSDSPVWGSGGVYQKAVVEPSCSE 291
DB 419 NDEINPIKALHLELQHSIIITINSQPOYNGIRSKDKIHGMPDGYVYQKHLEFRLPKT 478
QY 292 KIDALVDKCKRFTSLTYAANKDSWMSN-VGQTDNAVATMGVPAKEIIQPTIVDPVSF 350
DB 479 KLPKRLDITKNNFLLTYFAIDSGDLISNHPDNKSNKAWGIFPREILOPTEIVEKISF 538
QY 351 NYWKDEAFETISGRWASLYPEDEA-SRKLVEVGSFSLVSLNDYINGD 400
DB 539 LMKKEFFYIILNEMKLNKKYDKRHSQFIQSLDQCLVINYINDYISPD 589

RESULT 5
T15423
hypothetical protein C06A8.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15423
R:Leinbach, D.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C06A8.
A:Reference number: Z18348
A:Accession: T15423

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-615 <LE>
A:Cross-references: EMBL:U39849; NID:91055041; PID:91055045; PIDN:AAA81048.1; CESP:C0
A:Genetics:
A:Gene: CESP:C06A8.1
A:Introns: 28/2; 53/2; 177/3; 375/3; 513/3; 565/2

Query Match 32.5%; Score 711; DB 2; Length 615;
Best Local Similarity 38.4%; Pred. No. 8.3e-49;
Matches 158; Conservative 63; Mismatches 126; Indels 64; Gaps 10;

1 ARVYOLFVDFDIFLKVNDRCRIGITCPVPGIMPINNYKGFIMTGFCKTKIPADIM 60
DB 235 ANFVITQLEFEKEFEKVRRCREIGITQPIIPGIMPINGESIKRIALSGLEIPQIHL 294
QY 61 AALPEIKDNEKAVKAYGIIHGTCEKCKILAHG-IKTLHLYTLMEKSALAILMNLGLIEE 119
DB 295 DDEPKHDDDAVQKYGTERCIEKCRRLDNGTAPSIHLYTMNRESIIEILKSLGL-- 351
QY 120 SKVRSLLPMRRPANYRVEDVRPIFMANR--PKSYISKTIGMDQYPRGMRGDSQNP 176
DB 352 -----WK-----LEGDRVFPWKNSQHPHOLESTV----- 377
QY 177 GALSQYQMRPRANDKRLVEEAVPLKSVEDYERFRLYCLG-----KLASNWSELD 229
DB 378 -----RNADRLA-MEGANISFEDVRKRVFINTYQAPNADGVKIVLFWTAE 425
QY 230 -GLQPEYKILNQLEKINTKGFILINSOPAVNGEKSDSPVWGSGGVYQKAVVEPFC 288
DB 426 TGVQPEISLISQILWNCNENILTVNSQPSVNGAPSTPLVWGKRGYCYQKAYLECFM 485
QY 289 SNEKIDALVDKCKR--TSLTYAANKDSW-KSNQGDVNAVATMGVPAKEIIQPTIV 345
DB 486 TAEISDKLIQIERFPRVRYHAIKNDSTFDKTNSETTIPAVITGVYFSGEIIQPTIV 545
QY 346 DPVSFVNMKDEAFETISGRWASLYPEDEASRKLVEEVGSFSLVSLNDY 396
DB 546 DPLSRANWDEAYQWMAQMDYFPKESKSGYIKAVHDEFRLVTLVNDYF 596

RESULT 6
T38659
methylenetetrahydrofolate reductase 2 - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000
C:Accession: T38659
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21804
A:Accession: T38659
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-641 <MD>
A:Cross-references: EMBL:AL109739; PIDN:CA822273.1; GSPDB:GN00066; SPDB:SPAC343.10
A:Experimental source: strain 972h-; cosmid c343
C:Genetics:
A:Gene: SPDB:SPAC343.10
A:Map position: 1

Query Match 28.8%; Score 629; DB 2; Length 641;
Best Local Similarity 32.2%; Pred. No. 3.2e-42;
Matches 147; Conservative 80; Mismatches 161; Indels 68; Gaps 13;

1 ARVYOLFVDFDIFLKVNDRCRIGITCPVPGIMPINNYKGFIMTGFCKTKIP 56
DB 184 ADRITQIYFEARAIKENVVRNHSNALRNPIIPALIMPISQYSGLRMRRLCGCVP 243
QY 57 ADIMALEPIKNEKAVKAYGIIHGTCEKCKIL-----AHGKITLHLYTLMEKSALAI 110
DB 244 SSLMQRNLNAKPPDEALKNIGVEHITVMIKIMDNQGVHVG---FHFCTLERSVALI 300
QY 111 LKNLGLI-----EESKVSRS-----LPMRRP-----ANVR 136

```

Db      301 LKNSGLIFKRWKQYSEMEDEKLTTRKRLSLDEPAELHNOYVPSQPPVADKSSNLFV 360
OY      137 VAE-----DVPPIFWANRPKSYISRTIGMDQPHGWMGSCSPGALSDV-- 182
Db      361 TSKQSSVSGHNDNLTEAP-FTVSESGSVGLGRQANDDFITNGFQDPRAPELGEIDGCP 419
OY      183 QPMRPARDKALVEEMAVPLKSVEDIYERFLYCLGLKNSPWEISLQIPETKILINQL 242
Db      420 TLHPFPEALKL--WGYPVDE-SDITSLFQKHIMSDISAPMID-EPEVEETKIARYL 474
OY      243 EKINTKGLTINSOPAVNGEKSDSPYMGCGPGGYOKAYVEFCKSEKEDALVDKCD 302
Db      475 LKLNNSMWTGSPAPVANGASADPPVGMGKGRVQKAFVCEFNKDLKDFITKWD 534
OY      303 RSLTYMAYNNKDGSMKSVGQTDVNAVTVMPAKELIPTIYDPSFMVWDEAFETWS 362
Db      535 NQVITYYAGNNKSEFLTNAPKDGASATVTCGYRGRELIGSTIIAEVSFKAMLSSEFYWG 594
OY      363 RGMASLYPEDEASRKLYVEVGSHFLVSLVDNYIN 398
Db      595 E-WANLYSKNTPSRKLENCINDRMVLTVIHDEM 629

```

RESULT 7
 S63459
 methylentetrahydrofolate reductase (MADPH2) (EC 1.5.1.20) MET12 - yeast (Saccharomyces
 N:Alternate names: protein LP8C; protein YPL023c
 C:Species: saccharomyces cerevisiae
 C>Date: 16-May-1996 #sequence_revision 12-Jul-1996 #text_change 03-Jun-2002
 C:Accession: S63459
 R:Wang, Y.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; VC
 submitted to the EMBL Data Library, September 1995
 A:Reference number: S63452
 A:Accession: S63459
 A:Molecule type: DNA
 A:Residues: 1-657 <MAN>
 A:Cross-references: EMBL:U036624; NID:g1276642; PIDD:AA68164.1; PID:g1039454; GSPDB:GN00
 C:Genetics:
 A:Gene: SGD:MET12; MIPS:YPL023c
 A:Cross-references: SGD:S0005944; MIPS:YPL023c
 A:Map position: 16L
 C:Keywords: oxidoreductase

```

Query Match      26.5%: Score 578.5; DB 2: Length 657;
Best Local Similarity 31.5%: Pred. No. 3.6e-38;
Matches 151; Conservative 80; Mismatches 155; Indels 93; Gaps 16;

OY      1 ARVYVQLFPTDIFLKVNDRCQ-IGITCPYPGIMPINNYKGFIRMTGFCFKTIPADI 59
Db      183 ADFVITQLFVDEKFLTEMLFRRIISQDPLPEPGIMPINSYLLFHRRAKLSHASIPAI 242
OY      60 MALEP-IKDNEAVKAYGIHLGTEMCKI---LANGIKTLHLTYLNMKSALATIM--- 112
Db      243 LSHFPELIGSDNNAKSGIVDILLELQELIYORTSGRIKGFHYTLNLEKALQIVQSP 302
OY      113 -----NLGLI-----EEKVSRSLPMRRPANNV 134
Db      303 VLSHYNSESSEGEDELTSGEISIEVNPLEDADGDIVLDDSSNEETVANRRK--RRHSL 360
OY      135 FRVKEDVREIFWANR-----PKSYIS-----RTIGMDQYPHG 166
Db      361 ----SAKLIF--NRAIVTEKGLRYNNENSGMPSKALISIKGHGTLGRDATWDEFPNG 414
OY      167 RMGDGDNPSYGLSDYQPMR--RAROKLVEEMAVPLKSVEDIYERFLYCLGLRBNP 224
Db      415 RFEDGSRSPYGEIDGY--GSIIVSKSKALELMGP-KTIGLKLQIFIKYLEGSTDAIP 470
OY      225 WSELQLOPETKIINEOLEKINTKGLTINSOPAVNGEKSDSPYMGCGPGGYOKAYV 284
Db      471 WSDL-GLSERFALIQEELIQLNRYGVLTLASOPATNATLSSDKIFGWMGPAKRLYQKAFV 529
OY      285 EFFCKSEKEDALVDKCD---RTSLTYMAYNNKDGSMKSVNGQTDVNAVTVMPAKELI 341

```

```

Db      530 EMFIRQOEMTTLKPKLDHYGRKRFSSYAGDSSGSEFTMLDGHSSVYTWGVFPNRPVK 589
OY      342 PIIVYVSEFNWKKGEAFETWSRGMASLYPEDEASRKLYVEVGSHFLVSLVDNYIND 400
Db      590 TTIIEESFKARNDENISME-WAKLFPRKNTANLLLVHKDCVLSYVHHDFKED 647

```

RESULT 8
 S46454
 5,10-methylentetrahydrofolate reductase (FADH2) (EC 1.7.99.5) - human
 C:Species: Homo sapiens (man)
 C>Date: 15-Jul-1995 #sequence_revision 17-Nov-1995 #text_change 07-May-1999
 C:Accession: S46454
 R:Goette, P.; Sumner, J.S.; Miles, R.; Duncan, A.M.V.; Rosenblatt, D.S.; Matthews, R
 Nature Genet. 7, 195-200, 1994
 A:Title: Human methylentetrahydrofolate reductase: isolation of cDNA, mapping and mu
 A:Reference number: S46454; M01D:95004587; PMID:7920641
 A:Accession: S46454
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <GOY>
 C:Keywords: FAD; oxidoreductase

```

Query Match      22.2%: Score 486; DB 2: Length 416;
Best Local Similarity 46.4%: Pred. No. 4.7e-31;
Matches 90; Conservative 35; Mismatches 67; Indels 2; Gaps 2;

OY      1 ARVYVQLFPTDIFLKVNDRCQIGITCPYPGIMPINNYKGFIRMTGFCFKTIPADI 60
Db      223 ADFITQLFEPADFFPRVACADMGTICPIYGPIDIGYSLQVLKLSLELVEPEIK 282
OY      61 AALPEIKNEAVKAYGIHLGTEMCKIANG-IKTHLYTLNMKSALATIMNLGLIE 119
Db      283 DYIEPIKNDAAINNYGIELAVLSCQELASGLVPELHRYTLNRMATVEVLKRLGMWTE 342
OY      120 SKVSSSLPMRRPANNVFRVKEDVPRIPMANPKSYISRTIGMDQPHGWMGSCSPGAL 179
Db      343 DP-RKPLWALSAPHRKREEDVPRITWASRPKSYIRTEWDEFPNGMSSSPAREGL 401
OY      180 SDYOFMRPRARDKK 193
Db      402 KDYLEFLYLSKSPK 415

```

RESULT 9
 T42227
 methylentetrahydrofolate reductase homolog - fission yeast (Schizosaccharomyces pombe
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T42227
 R:Yoshidoka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: 217923; M01D:98162722; PMID:9501991
 A:Accession: T42227
 A:Status: preliminary; translated from CB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-566 <YOS>
 A:Cross-references: EMBL:D89118; NID:g1749443; PIDD:BA13780.1; PID:g1749444
 A:Experimental source: strain PR745

```

Query Match      19.3%: Score 421.5; DB 2: Length 566;
Best Local Similarity 30.6%: Pred. No. 1e-25;
Matches 110; Conservative 56; Mismatches 127; Indels 67; Gaps 12;

OY      1 ARVYVQLFPTDIFLKVNDRCQIGITC---PIVPGIMPINNYKGFIRMTGFCFKTIP 56
Db      174 ADFITQLFEPDEFIFEDFVRNHSNALRNPIIPAIPIOSYGLKRWTRLCGGSVP 233
OY      57 ADIMALPIDDNEAVKAYGIHLGTEMCKI-----AHGIKTLHLTYLNMKSALAI 110
Db      234 SSIQRLUAAKPRDDEALKNIGVEHYIDMKIMDNQAVRHG---FHCITLNRSAVALI 290

```

```

QY 111 IANLIGIT-----EESKVSRS-----LPMRR-----ANVR 136
      | | | | |
Db 291 LKNSGLITRKMKOYSEKEMDEKLTTRKRSLDEPPELHANOYVPSQDQVADKSSNLEV 350
      | | | | |
QY 137 VKE-----DVREIFWANRPKGYISRTIGMDQYPHRGWSDCNPSYCALSDY-- 182
      | | | | |
Db 351 TSKQSSVSGHKDNLTEAP-FVSESGSVLGRQANMDDFTNGRRGDPDRSPAYGEIDGYCP 409
      | | | | |
QY 183 QFMRRPRRDKKLYEENAVPLKSEDIYERFRLYCLGKLRSPMSLDDGLOPEKTIINEOJ 242
      | | | | |
Db 410 TLHPPESEALKL---WGPEVDE-SDITSLFQKHMSDISVLPWHD-BPEVEKTRITAKXL 464
      | | | | |
QY 243 EKINTKGFLINSGPVPNNGEKSDSPYWGGEVGVYOKAVPEFCSKEKLDALVYKXD 302
      | | | | |
Db 465 LKLNKMSMWTVGSOQPVNNGCAPSADPEVFGMGPKRGVAFDOKAVCEVFNGLKLPFTTKMHD 524

```

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
A.; Antunes, da Silva, A.C.R.; da Silva F.R.; da Silva, A.M.; Silva Jr., M.A.; da Silva
M.; Tshahko, M.H.; Valada, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A.; Reference number: A59328
A; Contents: annotation
C; Genetics:
A; Gene: XP121
C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

Query Match 9.8%; Score 215; DB 2; Length 275;
Best Local Similarity 33.9%; Pred. No. 1.2e-09;

Matches 39; Conservative 30; Mismatches 46; Indels 0; Gaps 0;

QY 1 ARVITQLEFYDIDFLKFNDCRQIGTCPIYVGIMPINNKGFTRMTGFCCKTKIPADIM 60
DB 160 ADAAITQYFNPNFHFVDVAVRGVITPIYAGVPLSNEDKLNHFSGCCGAEIPRWIT 219
QY 61 AALEPIKNEAVKAYGIGHGTEMCKKILAHGIKTLHLTYLNMKSALAILMNLGL 115
DB 220 KKMAYGDDTKSIRAFGADVATLCERLIAGAGPLHFTLNLAKPSTQVLRIG 274

RESULT 14

5,10-methylenetetrahydrofolate reductase - Streptomyces coelicolor

C; Species: Streptomyces coelicolor

C; Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000

C; Accession: T34973

R; Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999

A; Reference number: Z21563

A; Accession: T34973

A; Status: preliminary; translated from GB/EMBL/DBJ

A; Molecule type: DNA

A; Residues: 1307 <SN>

A; Cross-references: EMBL:AL109663; PIDN:CAB52012.1; GSPDB:GN00070; SCOEDB:SC4A10.36C

A; Experimental source: strain A3(2)

C; Genetics:

A; Gene: metF; SCOEDB:SC4A10.36C

C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

Query Match 8.8%; Score 193; DB 2; Length 307;
Best Local Similarity 33.0%; Pred. No. 8.3e-08;
Matches 37; Conservative 22; Mismatches 53; Indels 0; Gaps 0;

QY 5 VTQLEFYDIDFLKFNDCRQIGTCPIYVGIMPINNKGFTRMTGFCCKTKIPADIMALE 64
DB 190 ITQMFQPDSDYRLRDVAAAGATPVIPVSVMLERLPLKLSMASFPALKEKITL 249

QY 65 PIKNEAVKAYGIGHGTEMCKKILAHGIKTLHLTYLNMKSALAILMNLGL 116
DB 250 TAKDDPAVARSIGIEFATFCARLLAEGVPGHFTLNNSTATLEIYENGL 301

RESULT 15

5,10-methylenetetrahydrofolate reductase (FADH2) (EC 1.7.99.5) - Haemophilus influenzae

C; Species: Haemophilus influenzae

C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C; Accession: H64123

R; Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Meldrum, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995

A; Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A; Reference number: A64000; MUID: 95350630; PMID: 7542800

A; Accession: H64123

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-292 <TIGR>

A; Cross-references: GB:U32823; GB:L42023; NID:g1574281; PIDN:AMC23094.1; PID:g1574284; T

C; Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

C; Keywords: methionine biosynthesis; oxidoreductase

Query Match 8.6%; Score 189; DB 1; Length 292;
Best Local Similarity 32.8%; Pred. No. 1.6e-07;
Matches 38; Conservative 23; Mismatches 55; Indels 0; Gaps 0;

QY 1 ARVITQLEFYDIDFLKFNDCRQIGTCPIYVGIMPINNKGFTRMTGFCCKTKIPADIM 60
DB 175 ANHVTIOFFDIENYLRFRDRCASIGIDTEIYVGLPVTNFKQLOKMASTNNKIPAMLV 234
QY 61 AALEPIKNEAVKAYGIGHGTEMCKKILAHGIKTLHLTYLNMKSALAILMNLGL 116
DB 235 KAYDGLNDPTTNRNLVAASVAMDKVILTSREGVNDPHFTLNRSELTATICHMIGV 290

Search completed: February 14, 2003, 23:10:48
Job time : 44 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 18:14:55 ; Search time 70 Seconds

(without alignments)
776.661 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187
Sequence: 1 ARVIVTQLFVYDIDFLKFN.....VSLVNDXINGDLFAVAFAD 408

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2187 | 100.0 | 408 | 21 | AAV44740 |
| 2 | 1834 | 83.9 | 554 | 21 | AA647726 |
| 3 | 1834 | 83.9 | 557 | 21 | AA647725 |
| 4 | 1834 | 83.9 | 566 | 21 | AA632205 |
| 5 | 1834 | 83.9 | 569 | 21 | AA632204 |
| 6 | 1834 | 83.9 | 594 | 21 | AA647724 |
| 7 | 1834 | 83.9 | 606 | 21 | AA632203 |
| 8 | 1760 | 76.1 | 391 | 21 | AA615450 |
| 9 | 1664 | 73.5 | 373 | 21 | AA615451 |
| 10 | 1608 | 73.5 | 362 | 21 | AA615452 |

| | | | | | | |
|----|-------|------|------|----|----------|---------------------|
| 11 | 1184 | 54.1 | 433 | 21 | AA636171 | Arabidopsis thalia |
| 12 | 1184 | 54.1 | 470 | 21 | AA636170 | Arabidopsis thalia |
| 13 | 1184 | 54.1 | 497 | 21 | AA636169 | Arabidopsis thalia |
| 14 | 979 | 44.8 | 656 | 17 | AA688358 | Human methylene-t |
| 15 | 979 | 44.8 | 656 | 17 | AA612607 | Human methylene-t |
| 16 | 979 | 44.8 | 660 | 17 | AA688359 | Human methylene-t |
| 17 | 979 | 44.8 | 660 | 21 | AAV96186 | Human methylene-t |
| 18 | 979 | 44.8 | 660 | 21 | AAE12606 | Human methylene-t |
| 19 | 979 | 44.8 | 660 | 23 | AAV75413 | Human methylene-t |
| 20 | 979 | 44.8 | 660 | 23 | AAV75421 | Human methylene-t |
| 21 | 979 | 44.8 | 660 | 23 | AAV75422 | Human methylene-t |
| 22 | 979 | 44.8 | 660 | 23 | AAV75423 | Human methylene-t |
| 23 | 979 | 44.8 | 660 | 23 | AAV75424 | Human methylene-t |
| 24 | 976 | 44.6 | 660 | 23 | AAV75425 | Human methylene-t |
| 25 | 973 | 44.5 | 660 | 23 | AAV75430 | Human methylene-t |
| 26 | 971 | 44.4 | 660 | 23 | AAV75431 | Human methylene-t |
| 27 | 969 | 44.3 | 660 | 23 | AAV75432 | Human methylene-t |
| 28 | 293 | 13.4 | 101 | 21 | AAV44741 | Wheat 5,10-methyl |
| 29 | 182 | 8.3 | 292 | 23 | AAV72976 | Lactococcus lactis |
| 30 | 163 | 7.5 | 283 | 23 | AA654590 | Corynebacterium gl |
| 31 | 158 | 7.2 | 294 | 22 | AA679970 | C. glutamicum prote |
| 32 | 158 | 7.2 | 349 | 22 | AA692125 | C. glutamicum melar |
| 33 | 157 | 7.2 | 349 | 22 | AA607695 | Streptomyces fradi |
| 34 | 155 | 7.1 | 349 | 22 | AA653361 | Human OREF protein |
| 35 | 138 | 6.3 | 305 | 23 | AAE13585 | Human OREF protein |
| 36 | 115 | 5.3 | 69 | 23 | ABP07534 | Novel human diapo |
| 37 | 106.5 | 4.9 | 97 | 23 | ABP11142 | Drosophila melanog |
| 38 | 103 | 4.7 | 1120 | 22 | ABG19663 | Vesicular stomatit |
| 39 | 99.5 | 4.5 | 619 | 22 | AB64250 | P. falciparum Proj |
| 40 | 97.5 | 4.5 | 422 | 22 | AA620281 | P. falciparum Proj |
| 41 | 97.5 | 4.5 | 2703 | 16 | AA670236 | Plasmodium var-7 |
| 42 | 97.5 | 4.5 | 2710 | 18 | AA622482 | Plasmodium var-7 |
| 43 | 97.5 | 4.5 | 2710 | 21 | AAV77904 | Plasmodium var-7 |
| 44 | 97.5 | 4.5 | 3060 | 21 | AAV22475 | Plasmodium var-7 |
| 45 | 97.5 | 4.5 | 3060 | 21 | AAV77905 | Plasmodium var-7 |

ALIGNMENTS

| | | |
|----------|---|-------------------------------------|
| RESULT 1 | AAV44740 | AAV44740 standard; Protein; 408 AA. |
| ID | AAV44740 | |
| XX | AAV44740; | |
| AC | AAV44740; | |
| XX | | |
| DT | 04-MAY-2000 (first entry) | |
| XX | | |
| DE | Soybean 5,10-methylenetetrahydrofolate reductase. | |
| XX | | |
| KW | 5,10-methylenetetrahydrofolate reductase; MTHFR; soybean; inhibitor; methionine synthesis; tetrahydrofolate metabolism enzyme; herbicide. | |
| XX | | |
| OS | Glycine max. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | Region | 1..3 |
| FT | /note= "No codons in the corresponding nucleotide sequence" | |
| FT | Region | 189..408 |
| FT | /note= "No codons in the corresponding nucleotide sequence" | |
| XX | | |
| PN | WO200004163-A1. | |
| XX | | |
| PD | 27-JAN-2000. | |
| XX | | |
| PF | 14-JUL-1999; | 99WO-US15916. |
| XX | | |
| PR | 15-JUL-1998; | 98US-0092869. |
| XX | | |
| PA | (DUPO) DU POINT DE MEMOURS & CO E. I. | |
| XX | | |

PI Falco SC, Fomodu LO:
XX
DR WPI: 2000-182429/16.
DR N-PSDB; AA250061.
XX
PT Novel tetrahydrofolate metabolism enzyme used to alter the level of
PT tetrahydrofolate metabolism in plants and seeds -
XX
PS Claim 6; Pages 32-34; 37pp; English.
XX
CC The present sequence is a 5,10-methylenetetrahydrofolate reductase
CC (MTHFR) encoded by cDNA clone sfil1.pk0017.d12. The clone was
CC isolated from a still cDNA library which was prepared using soybean
CC immature flowers. MTHFR (E.C. 1.7.99.5) plays a role in the
CC synthesis of methionine. The coding sequence of the enzyme is used in
CC the construction of a chimeric gene to alter the level of
CC tetrahydrofolate metabolism enzymes in plants. The enzyme may provide
CC target to facilitate design and/or identification of inhibitors that may
CC be useful as herbicides.
CC N.B. The present sequence is also stated to be encoded by the
CC ccol.pk0049.d4 clone in AA250060.
XX
SQ Sequence 408 AA:
Query Match 100.0%; Score 2187; DB 21; Length 408;
Best local similarity 100.0%; Pred. No. 1.2e-210;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAVITQLFYDTDFLEKFNVDCHQIGTCPIVPGIMPNYKGFIRMTGFKTKIPADIM 60
DB 1 AAVITQLFYDTDFLEKFNVDCHQIGTCPIVPGIMPNYKGFIRMTGFKTKIPADIM 60
QY 61 AALEPIKNEBNAVAGYGHGTECKKILAHGKITLHLYTLNMEKSALITMNGILEES 120
DB 61 AALEPIKNEBNAVAGYGHGTECKKILAHGKITLHLYTLNMEKSALITMNGILEES 120
QY 121 KVSRLPMRRPAPNFRVKEEDRPTEFMANRKYSTISITICMDQYPHGRMGDCSNPSYALS 180
DB 121 KVSRLPMRRPAPNFRVKEEDRPTEFMANRKYSTISITICMDQYPHGRMGDCSNPSYALS 180
QY 181 DVQFMKPRAROKKLYEEMAVPLKSVEDIYERFLYCLGRNPNSELDGLQPEKTIINE 240
DB 181 DVQFMKPRAROKKLYEEMAVPLKSVEDIYERFLYCLGRNPNSELDGLQPEKTIINE 240
QY 241 QLEKINTGFLTINSQPAVNGEKSDPTVGMGGPGGVYQKAYVEFCGSKRLDALYDKC 300
DB 241 QLEKINTGFLTINSQPAVNGEKSDPTVGMGGPGGVYQKAYVEFCGSKRLDALYDKC 300
QY 301 KQRTSLITMAVANKDGSMSKNVGGTDVNAVTVGVPAKETIIQPTIYDPVSFNWKDEAFET 360
DB 301 KQRTSLITMAVANKDGSMSKNVGGTDVNAVTVGVPAKETIIQPTIYDPVSFNWKDEAFET 360
QY 361 WSRGNASLYPEDEASRKLVEEVGSHFLVSLVNDYIINGDLFAVEADF 408
DB 361 WSRGNASLYPEDEASRKLVEEVGSHFLVSLVNDYIINGDLFAVEADF 408
RESULT 2
AA647726
ID AA647726 standard; Protein: 554 AA.
XX
AC AAG47726;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60184.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0125448.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134944.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0138452.
PR 17-JUN-1999; 99US-0138453.
PR 17-JUN-1999; 99US-0138492.
PR 18-JUN-1999; 99US-0138454.
PR 18-JUN-1999; 99US-0138455.
PR 18-JUN-1999; 99US-0138456.
PR 18-JUN-1999; 99US-0138457.
PR 18-JUN-1999; 99US-0138458.
PR 18-JUN-1999; 99US-0138459.
PR 18-JUN-1999; 99US-0138460.
PR 18-JUN-1999; 99US-0138461.
PR 18-JUN-1999; 99US-0138462.
PR 18-JUN-1999; 99US-0138463.
PR 18-JUN-1999; 99US-0138464.
PR 18-JUN-1999; 99US-0138465.
PR 21-JUN-1999; 99US-0138466.
PR 22-JUN-1999; 99US-0138467.
PR 23-JUN-1999; 99US-0138468.
PR 23-JUN-1999; 99US-0138469.
PR 24-JUN-1999; 99US-0138470.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.


```
RESULT 3
AAG47725
ID AAG47725 standard; Protein: 557 AA.
XX
AC AAG47725;
XX
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SPQ ID NO: 60183.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX
XX
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 24-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 26-JUN-1999; 99US-0140823.
XX 28-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 02-JUL-1999; 99US-0142154.
XX 06-JUL-1999; 99US-0142055.
XX 08-JUL-1999; 99US-0142390.
XX 09-JUL-1999; 99US-0142803.
XX 12-JUL-1999; 99US-0142920.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 05-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
XX 09-AUG-1999; 99US-0147935.
XX 10-AUG-1999; 99US-0148171.
XX 11-AUG-1999; 99US-0148319.
XX 12-AUG-1999; 99US-0148341.
XX 13-AUG-1999; 99US-0148565.
XX 13-AUG-1999; 99US-0148684.
XX 16-AUG-1999; 99US-0149368.
XX 17-AUG-1999; 99US-0149175.
XX 18-AUG-1999; 99US-0149426.
XX 20-AUG-1999; 99US-0149722.
```



```
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      83.9%: Score 1834: DB 21: Length 557;
Best Local Similarity 82.1%: Pred No. 57e-175;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;
```

```
QY 1 ARVYTOLEFDYDFILKFNVDCKQIGITCPVIGIMPINNKKGITRMGFCFKTKIPADIM 60
DB 150 ADLVYTOLEFDYDFILKFNVDCKQIGITCPVIGIMPINNKKGITRMGFCFKTKIPAEVM 209
QY 61 AALEPIKNEEAVAYGIHGTBMCKKILAHGKITLHLYTLNMEKSALATIMNLGLEES 120
DB 210 AALEPIKNEEAVAYGIHGTBMCKKMLAHGVKSLHLYTLNMEKSALATIMNLGMIDES 269
QY 121 KYSRSLPFRRRPANYFRKEDVRPIFWANRPKYSISTKTICWDYPIHGRNGDSCNFSYGALS 180
DB 270 KISRSLPFRRRPANYFRKEDVRPIFWANRPKYSISTKTICWDYPIHGRNGDSCNFSYGALS 329
```

```
QY 181 DYOPMRPRANDKLYEEMAVPLEKSVEDYERFERFYCLGKLSNPMSELDLOPETKINE 240
DB 330 DHQSRPRANDKRLQOEWVYPLKSVEDIOEKFKELCUGNLKSSPWSMDLOPETKINE 389
QY 241 QLEKINTKGEITINSOPAVNGEKSDSPYWGCGEGCIYYOKAYIEFCSKEKIDALYDC 300
DB 390 QLIKVNSKGFLTINSOPVNAERSDSPYWGCGEGCIYYOKAYIEFCSKEKIDAVEKC 449
QY 301 KDRTSLTYMAVNVKDSGMSKNVCGQTFVNAVTVGWVPAKEIIPITIVDPVSFNWKADEAFET 360
DB 450 KALPSTIYMAVNVKGEQWVNSTAQADVNAVTVGWVPAKEIIPITIVDPVSFNWKADEAFET 509
QY 361 WSRGMASLYPEDASRKLYEEVGGSHFLVSLYDNDYINGDIFAFAD 407
DB 510 WSRGMASLYPEDASRRLLEEVAKNSTYVSLVENDYINGDIFAFAD 556

RESULT 4
AAG32205
ID AAG32205 standard; Protein; 566 AA.
AC AAG32205;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38806.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135523.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138940.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.
Query Match 83.9%; Score 1834; DB 21; Length 566;
Best Local Similarity 82.1%; Pred. No. 5,8e-175;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;
QY 1 ARVIYTOLEFYDPTDFLEKLVNDYCRQIGITCPVPGIMPINNKGFIKIRYTGCKTKIPADIM 60
|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
DB 159 ADLIYTOLEFYDPTDFLEKLVNDYCRQIGITCPVPGIMPINNKGFIKIRYTGCKTKIPADIM 218
61 AALEPIKONEBAVAVYGIHLCGTECKKILAHGKFTLHLYLNMEKSAIATLMLGLIEES 120
DB 219 AALEPIKONEBAVAVYGIHLCGTECKKILAHGKFTLHLYLNMEKSAIATLMLGLIDES 278
QY 121 KVSRLPMPRPANRYVEDYRPFEMANRPKSYISRTIGMDQYPHGWDGSCNPSYGAIS 180
|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
DB 279 KISRLPMPRPANRYVEDYRPFEMANRPKSYISRTIGMDQYPHGWDGSCNPSYGAIS 338
QY 181 DYQMRPRARDDKLVEMAVPLKSVEDIYERFRLYCLGKLRSPWSELDGLQPETRIINE 240
DB 339 DHQFRPRARDDKLVEMAVPLKSVEDIYERFRLYCLGKLRSPWSELDGLQPETRIINE 398
QY 241 QLEKNTGCGFLTINSQAPNGEKSDSPYVGWGGVGVYQKAVVEPFCSEKLDALVDC 300
|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
DB 399 QLIKNSGCGFLTINSQAPNGEKSDSPYVGWGGVGVYQKAVVEPFCSEKLDALVDC 458
QY 301 KDRISLYTMAVNDGSKSNVGTQDVNAVTVGYFPAKEIIQPTIVDFVSFNKDEAFET 360
|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
DB 459 KALPITYMAVNDGSKSNVGTQDVNAVTVGYFPAKEIIQPTIVDFVSFNKDEAFET 518
QY 361 WSRGASLYPEDEASRKLVEEVGSHFVLSVNDYINGDLFVFEAD 407
|:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
DB 519 WSRGASLYPEDEASRKLVEEVGSHFVLSVNDYINGDLFVFEAD 565

RESULT 5
AAG32204
ID AAG32204 standard; Protein: 569 AA.
XX AAG32204;
AC AAG32204;
XX AAG32204;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38805.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125768.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.

| | | |
|----|-------------|--------------|
| PR | 21-JUL-1999 | 9905-0145088 |
| PR | 22-JUL-1999 | 9905-0145085 |
| PR | 22-JUL-1999 | 9905-0145087 |
| PR | 22-JUL-1999 | 9905-0145092 |
| PR | 22-JUL-1999 | 9905-0145158 |
| PR | 23-JUL-1999 | 9905-0145145 |
| PR | 23-JUL-1999 | 9905-0145218 |
| PR | 23-JUL-1999 | 9905-0145524 |
| PR | 26-JUL-1999 | 9905-0145576 |
| PR | 27-JUL-1999 | 9905-0145513 |
| PR | 27-JUL-1999 | 9905-0145518 |
| PR | 28-JUL-1999 | 9905-0145591 |
| PR | 02-AUG-1999 | 9905-0146386 |
| PR | 02-AUG-1999 | 9905-0146388 |
| PR | 02-AUG-1999 | 9905-0146389 |
| PR | 03-AUG-1999 | 9905-0147038 |
| PR | 04-AUG-1999 | 9905-0147204 |
| PR | 04-AUG-1999 | 9905-0147302 |
| PR | 05-AUG-1999 | 9905-0147760 |
| PR | 05-AUG-1999 | 9905-0147760 |
| PR | 06-AUG-1999 | 9905-0147303 |
| PR | 06-AUG-1999 | 9905-0147416 |
| PR | 09-AUG-1999 | 9905-0147493 |
| PR | 09-AUG-1999 | 9905-0147535 |
| PR | 10-AUG-1999 | 9905-0148171 |
| PR | 11-AUG-1999 | 9905-0148319 |
| PR | 12-AUG-1999 | 9905-0148341 |
| PR | 13-AUG-1999 | 9905-0148565 |
| PR | 13-AUG-1999 | 9905-0148684 |
| PR | 16-AUG-1999 | 9905-0149368 |
| PR | 17-AUG-1999 | 9905-0149191 |
| PR | 18-AUG-1999 | 9905-0149426 |
| PR | 20-AUG-1999 | 9905-0149622 |
| PR | 20-AUG-1999 | 9905-0149723 |
| PR | 20-AUG-1999 | 9905-0149929 |
| PR | 22-AUG-1999 | 9905-0149930 |
| PR | 23-AUG-1999 | 9905-0149930 |
| PR | 23-AUG-1999 | 9905-0149930 |
| PR | 26-AUG-1999 | 9905-0150566 |
| PR | 27-AUG-1999 | 9905-0150884 |
| PR | 27-AUG-1999 | 9905-0151065 |
| PR | 27-AUG-1999 | 9905-0151066 |
| PR | 30-AUG-1999 | 9905-0151303 |
| PR | 31-AUG-1999 | 9905-0151338 |
| PR | 01-SEP-1999 | 9905-0151340 |
| PR | 07-SEP-1999 | 9905-0152367 |
| PR | 10-SEP-1999 | 9905-0153070 |
| PR | 13-SEP-1999 | 9905-0153158 |
| PR | 15-SEP-1999 | 9905-0154079 |
| PR | 16-SEP-1999 | 9905-0154739 |
| PR | 20-SEP-1999 | 9905-0155479 |
| PR | 22-SEP-1999 | 9905-0155139 |
| PR | 23-SEP-1999 | 9905-0155866 |
| PR | 24-SEP-1999 | 9905-0156369 |
| PR | 28-SEP-1999 | 9905-0156458 |
| PR | 29-SEP-1999 | 9905-0156596 |
| PR | 04-OCT-1999 | 9905-0157117 |
| PR | 05-OCT-1999 | 9905-0157753 |
| PR | 06-OCT-1999 | 9905-0157865 |
| PR | 07-OCT-1999 | 9905-0158029 |
| PR | 08-OCT-1999 | 9905-0158232 |
| PR | 12-OCT-1999 | 9905-0158369 |
| PR | 13-OCT-1999 | 9905-0159293 |
| PR | 13-OCT-1999 | 9905-0159294 |
| PR | 14-OCT-1999 | 9905-0159329 |
| PR | 14-OCT-1999 | 9905-0159330 |
| PR | 14-OCT-1999 | 9905-0159637 |
| PR | 14-OCT-1999 | 9905-0159637 |
| PR | 18-OCT-1999 | 9905-0159684 |
| PR | 21-OCT-1999 | 9905-0160741 |

| | | |
|----|-------------|--------------|
| PR | 21-OCT-1999 | 9905-0160767 |
| PR | 21-OCT-1999 | 9905-0160768 |
| PR | 21-OCT-1999 | 9905-0160770 |
| PR | 21-OCT-1999 | 9905-0160814 |
| PR | 21-OCT-1999 | 9905-0160815 |
| PR | 21-OCT-1999 | 9905-0160980 |
| PR | 22-OCT-1999 | 9905-0160981 |
| PR | 22-OCT-1999 | 9905-0160984 |
| PR | 22-OCT-1999 | 9905-0160989 |
| PR | 25-OCT-1999 | 9905-0161404 |
| PR | 25-OCT-1999 | 9905-0161405 |
| PR | 25-OCT-1999 | 9905-0161406 |
| PR | 25-OCT-1999 | 9905-0161350 |
| PR | 25-OCT-1999 | 9905-0161359 |
| PR | 26-OCT-1999 | 9905-0161361 |
| PR | 26-OCT-1999 | 9905-0161320 |
| PR | 28-OCT-1999 | 9905-0161920 |
| PR | 28-OCT-1999 | 9905-0161943 |
| PR | 29-OCT-1999 | 9905-0162142 |

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 83.9%; | Score 1834; | DB 21; | Length 569; |
| Best Local Similarity | 82.1%; | Pred. No. 5.9e-175; | | |
| Matches 334; | Conservative 37; | Mismatches 36; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|---|-----|
| QY | 1 | ANAYVQFLPFDIDDIFLAKFPANCROGIGTCPIVBPIMPINNYKGFIMNGFCKTKIPADIM | 60 |
| Db | 162 | ADLLVYQFLPFDIDDIFLAKFPANCROGIGTCPIVBPIMPINNYREFLMNGFCKTKIPVEVM | 221 |
| QY | 61 | AALEPIKDNEAVKAGVGHILGTEMCKKILLAGIKITLHLYLNNEKSALAILNNGLIEES | 120 |
| Db | 222 | AALEPIKDNEAVKAGVGHILGTEMCKKILLAGVKSLSHLTYLNNEKSALAILNNGMIDES | 281 |
| QY | 121 | KVSHSLPWRBPANVFKEKEDVBPFIWMANRPKSYISRTIGMDQYPHGRWGDSCNPSYGALS | 180 |
| Db | 282 | KISPSLPWRBPANVFETKEDVBPFIWMANRPKSYISRTIGMEDPQGRWGDSCNPSYGALS | 341 |
| QY | 181 | DYQMPRRADKKLYBEAVNPLKSYVEDIYEFKFLICLGKRSNPSSELDGLOPETKIINE | 240 |
| Db | 342 | DHQSRRPADKKLYBEAVNPLKSYVEDIOEFKFLICLGKRSNPSSELDGLOPETRIINE | 401 |
| QY | 241 | QLEIKNGKGLTINSOPAVNGEKSSPIVGGGGGGVYOKAVEPFCSEKELDALVDKC | 300 |
| Db | 402 | QLIKVNSKGLTINSOPAVNBERSDSPVGGGGGVYVOKAYLEPFCSEKELDALVEKC | 461 |
| QY | 301 | KDRTSLTYMAVKNKDGSMKSNVGGQTDVNAVTVGWGFPAKEIIQPTIVDPVSFNWNRDEAFEI | 360 |
| Db | 462 | KALPSIITYMAVNGEQWVSNTQADAVNAVTVGWGFPAKEIIQPTIVDPASFNWNRDEAFET | 521 |
| QY | 361 | WSRKMASLIPDEDBASKLYEEVGGSHFLYSLVNDYDINDGLAVRAVD | 407 |
| Db | 522 | WSRKMANLYPEADPSRNLLLEKKSYSLSLVENDYDINDGIVRAVD | 568 |

| | |
|----------|--|
| RESULT | 6 |
| AA647724 | |
| ID | AA647724 standard; Protein: 594 AA. |
| XX | |
| AC | AA647724; |
| XX | |
| DT | 18-OCT-2000 (first entry) |
| XX | |
| DE | Arabidopsis thaliana protein fragment SEQ ID NO: 60182. |
| XX | |
| KW | Protein identification; signal transduction pathway; metabolic pathway; |
| KW | hybridisation assay; genetic mapping; gene expression control; promoter; |
| KW | termination sequence. |
| XX | |
| OS | Arabidopsis thaliana. |
| XX | |
| PN | EPI033405-A2. |
| XX | |
| PD | 06-SEP-2000. |
| XX | |
| PF | 25-FEB-2000; 2000EP-0301439. |
| XX | |

| | | | | | |
|----|--------------|---------------|----|--------------|---------------|
| XX | 25-FEB-1999; | 99US-0121825; | PR | 12-JUL-1999; | 99US-0142977; |
| PR | 05-MAR-1999; | 99US-0123180; | PR | 11-JUL-1999; | 99US-0143542; |
| PR | 09-MAR-1999; | 99US-0123548; | PR | 14-JUL-1999; | 99US-0143624; |
| PR | 23-MAR-1999; | 99US-0125788; | PR | 15-JUL-1999; | 99US-0144005; |
| PR | 25-MAR-1999; | 99US-0126264; | PR | 16-JUL-1999; | 99US-0144085; |
| PR | 29-MAR-1999; | 99US-0126785; | PR | 19-JUL-1999; | 99US-0144086; |
| PR | 01-APR-1999; | 99US-0127462; | PR | 19-JUL-1999; | 99US-0144325; |
| PR | 06-APR-1999; | 99US-0128234; | PR | 19-JUL-1999; | 99US-0144331; |
| PR | 08-APR-1999; | 99US-0128714; | PR | 19-JUL-1999; | 99US-0144332; |
| PR | 16-APR-1999; | 99US-0129845; | PR | 19-JUL-1999; | 99US-0144333; |
| PR | 19-APR-1999; | 99US-0130077; | PR | 19-JUL-1999; | 99US-0144334; |
| PR | 21-APR-1999; | 99US-0130449; | PR | 19-JUL-1999; | 99US-0144335; |
| PR | 23-APR-1999; | 99US-0130510; | PR | 20-JUL-1999; | 99US-0144352; |
| PR | 23-APR-1999; | 99US-0130891; | PR | 20-JUL-1999; | 99US-0144362; |
| PR | 28-APR-1999; | 99US-0131449; | PR | 20-JUL-1999; | 99US-0144484; |
| PR | 30-APR-1999; | 99US-0132048; | PR | 21-JUL-1999; | 99US-0144814; |
| PR | 30-APR-1999; | 99US-0132407; | PR | 21-JUL-1999; | 99US-0145086; |
| PR | 04-MAY-1999; | 99US-0132484; | PR | 21-JUL-1999; | 99US-0145088; |
| PR | 05-MAY-1999; | 99US-0132485; | PR | 22-JUL-1999; | 99US-0145085; |
| PR | 06-MAY-1999; | 99US-0132486; | PR | 22-JUL-1999; | 99US-0145087; |
| PR | 07-MAY-1999; | 99US-0132487; | PR | 22-JUL-1999; | 99US-0145089; |
| PR | 11-MAY-1999; | 99US-0132863; | PR | 23-JUL-1999; | 99US-0145192; |
| PR | 14-MAY-1999; | 99US-0134256; | PR | 23-JUL-1999; | 99US-0145145; |
| PR | 14-MAY-1999; | 99US-0134218; | PR | 23-JUL-1999; | 99US-0145218; |
| PR | 14-MAY-1999; | 99US-0134219; | PR | 23-JUL-1999; | 99US-0145324; |
| PR | 14-MAY-1999; | 99US-0134221; | PR | 26-JUL-1999; | 99US-0145276; |
| PR | 14-MAY-1999; | 99US-0134370; | PR | 27-JUL-1999; | 99US-0145513; |
| PR | 18-MAY-1999; | 99US-0134768; | PR | 27-JUL-1999; | 99US-0145518; |
| PR | 18-MAY-1999; | 99US-0134941; | PR | 27-JUL-1999; | 99US-0145519; |
| PR | 20-MAY-1999; | 99US-0135124; | PR | 28-JUL-1999; | 99US-0145521; |
| PR | 21-MAY-1999; | 99US-0135353; | PR | 02-AUG-1999; | 99US-0146386; |
| PR | 24-MAY-1999; | 99US-0135629; | PR | 02-AUG-1999; | 99US-0146388; |
| PR | 25-MAY-1999; | 99US-0136021; | PR | 02-AUG-1999; | 99US-0146389; |
| PR | 27-MAY-1999; | 99US-0136392; | PR | 04-AUG-1999; | 99US-0147204; |
| PR | 28-MAY-1999; | 99US-0136782; | PR | 04-AUG-1999; | 99US-0147204; |
| PR | 01-JUN-1999; | 99US-0137222; | PR | 05-AUG-1999; | 99US-0147192; |
| PR | 03-JUN-1999; | 99US-0137502; | PR | 05-AUG-1999; | 99US-0147260; |
| PR | 04-JUN-1999; | 99US-0137502; | PR | 06-AUG-1999; | 99US-0147303; |
| PR | 07-JUN-1999; | 99US-0137724; | PR | 06-AUG-1999; | 99US-0147416; |
| PR | 08-JUN-1999; | 99US-0138094; | PR | 09-AUG-1999; | 99US-0147493; |
| PR | 10-JUN-1999; | 99US-0138540; | PR | 09-AUG-1999; | 99US-0147935; |
| PR | 10-JUN-1999; | 99US-0138847; | PR | 10-AUG-1999; | 99US-0148171; |
| PR | 14-JUN-1999; | 99US-0139119; | PR | 11-AUG-1999; | 99US-0148319; |
| PR | 16-JUN-1999; | 99US-0139452; | PR | 12-AUG-1999; | 99US-0148341; |
| PR | 16-JUN-1999; | 99US-0139453; | PR | 13-AUG-1999; | 99US-0148365; |
| PR | 17-JUN-1999; | 99US-0139492; | PR | 13-AUG-1999; | 99US-0148684; |
| PR | 18-JUN-1999; | 99US-0139454; | PR | 16-AUG-1999; | 99US-0149368; |
| PR | 18-JUN-1999; | 99US-0139455; | PR | 17-AUG-1999; | 99US-0149175; |
| PR | 18-JUN-1999; | 99US-0139456; | PR | 18-AUG-1999; | 99US-0149426; |
| PR | 18-JUN-1999; | 99US-0139457; | PR | 20-AUG-1999; | 99US-0149722; |
| PR | 18-JUN-1999; | 99US-0139458; | PR | 20-AUG-1999; | 99US-0149723; |
| PR | 18-JUN-1999; | 99US-0139459; | PR | 20-AUG-1999; | 99US-0149929; |
| PR | 18-JUN-1999; | 99US-0139460; | PR | 23-AUG-1999; | 99US-0149902; |
| PR | 18-JUN-1999; | 99US-0139461; | PR | 23-AUG-1999 | |

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160766.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 83.9%; Score 1834; DB 21; Length 594;
Best Local Similarity 82.1%; Pred. No. 6,3e-175;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

QY 1 ARVIYTOLEFYDIDIFLKFVNDRCROIGITGPYIPGIMPINNKYGFIRMGFECKTKIPADIM 60
DB 187 ADLIVTOLFEDYDIDIFLKFVNDRCROIGISCPYIPGIMPINNKYGFIRMGFECKTKIPADIM 246
QY 61 AALEPIKDNEEAVKAYGIHIGTEMCKKILAHGKTLHYTLNMEKSALAILNGLIEES 120
DB 247 AALEPIKDNEEAVKAYGIHIGTEMCKKILAHGKSHLYTLNMEKSALAILNGLIMIDES 306
QY 121 KVSRLPWRPAPVFRVKEEDVRPIFWANRPKSYISTIGWDYPHGRMGDCSNPSYGALS 180
DB 307 KISRLPWRPAPVFRVKEEDVRPIFWANRPKSYISTRTKGMEDFPOGRMGDSRSASYGALS 366
QY 181 DYQMRPRARDKLVEMAVPLKSVEDIERFRLYCLGKRSNPMSELDGLQPEKRIINE 240
DB 367 DHQSRPRARDKLVEMAVPLKSVEDIQKFEELCLGNKSPSELDGLQPEKRIINE 426
QY 241 QLEKINTKGFLLINSQPAVNGEKSDSPTVWGPGGYVYOKAYVEFFCSKEKLDALVDC 300
DB 427 QLIKVNKSGFLTINSQPSVAHERSDSPTVWGPGGYVYOKAYLEFFCSKEKLDANAVEKC 486
QY 301 KDRTSLTYMAVNVKSGSKSVNGQTDVNAVTVGWGFPKAKETIOPITVDPVAFNWKDEAFPI 360
DB 487 KALPSITYMAVNVKSGEOWVSTADADVNAVTVGWGFPKAKETIOPITVDPVAFNWKDEAFET 546
QY 361 WSRGWASLYPEDEASRKLVESGSHFLVSLVNDYINGDLFAVFAD 407
DB 547 WSRGWASLYPEADSPSRNLLEEVKNSTYLSLVLENDIINGDIFAVFAD 593

RESULT 7
AAG32203
ID AAG32203 standard; Protein; 606 AA.
XX

AC AAG32203;
XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 38804.
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38804.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

| | | |
|----|--------------|---------------|
| PR | 06-AUG-1999; | 990S-0147416. |
| PR | 09-AUG-1999; | 990S-0147493. |
| PR | 09-AUG-1999; | 990S-0147935. |
| PR | 10-AUG-1999; | 990S-0148171. |
| PR | 11-AUG-1999; | 990S-0148319. |
| PR | 12-AUG-1999; | 990S-0148341. |
| PR | 13-AUG-1999; | 990S-0148365. |
| PR | 13-AUG-1999; | 990S-0148684. |
| PR | 16-AUG-1999; | 990S-0149368. |
| PR | 17-AUG-1999; | 990S-0149175. |
| PR | 18-AUG-1999; | 990S-0149426. |
| PR | 20-AUG-1999; | 990S-0149722. |
| PR | 20-AUG-1999; | 990S-0149723. |
| PR | 20-AUG-1999; | 990S-0149929. |
| PR | 23-AUG-1999; | 990S-0149902. |
| PR | 25-AUG-1999; | 990S-0149930. |
| PR | 26-AUG-1999; | 990S-0150066. |
| PR | 27-AUG-1999; | 990S-0150084. |
| PR | 27-AUG-1999; | 990S-0151065. |
| PR | 27-AUG-1999; | 990S-0151066. |
| PR | 30-AUG-1999; | 990S-0151080. |
| PR | 31-AUG-1999; | 990S-0151303. |
| PR | 01-SEP-1999; | 990S-0151438. |
| PR | 07-SEP-1999; | 990S-0151930. |
| PR | 10-SEP-1999; | 990S-0152363. |
| PR | 13-SEP-1999; | 990S-0153070. |
| PR | 15-SEP-1999; | 990S-0153758. |
| PR | 16-SEP-1999; | 990S-0154018. |
| PR | 20-SEP-1999; | 990S-0154039. |
| PR | 22-SEP-1999; | 990S-0154479. |
| PR | 23-SEP-1999; | 990S-0155139. |
| PR | 24-SEP-1999; | 990S-0155486. |
| PR | 28-SEP-1999; | 990S-0155659. |
| PR | 29-SEP-1999; | 990S-0156458. |
| PR | 04-OCT-1999; | 990S-0156596. |
| PR | 05-OCT-1999; | 990S-0157117. |
| PR | 06-OCT-1999; | 990S-0157753. |
| PR | 07-OCT-1999; | 990S-0157865. |
| PR | 08-OCT-1999; | 990S-0158029. |
| PR | 12-OCT-1999; | 990S-0158232. |
| PR | 13-OCT-1999; | 990S-0158369. |
| PR | 13-OCT-1999; | 990S-0159293. |
| PR | 13-OCT-1999; | 990S-0159294. |
| PR | 13-OCT-1999; | 990S-0159295. |
| PR | 14-OCT-1999; | 990S-0159325. |
| PR | 14-OCT-1999; | 990S-0159329. |
| PR | 14-OCT-1999; | 990S-0159330. |
| PR | 14-OCT-1999; | 990S-0159331. |
| PR | 14-OCT-1999; | 990S-0159637. |
| PR | 18-OCT-1999; | 990S-0159638. |
| PR | 21-OCT-1999; | 990S-0159680. |
| PR | 21-OCT-1999; | 990S-0160741. |
| PR | 21-OCT-1999; | 990S-0160767. |
| PR | 21-OCT-1999; | 990S-0160768. |
| PR | 21-OCT-1999; | 990S-0160770. |
| PR | 21-OCT-1999; | 990S-0160814. |
| PR | 21-OCT-1999; | 990S-0160815. |
| PR | 22-OCT-1999; | 990S-0160980. |
| PR | 22-OCT-1999; | 990S-0160981. |
| PR | 22-OCT-1999; | 990S-0160982. |
| PR | 22-OCT-1999; | 990S-0160983. |
| PR | 25-OCT-1999; | 990S-0161404. |
| PR | 25-OCT-1999; | 990S-0161405. |
| PR | 25-OCT-1999; | 990S-0161406. |
| PR | 26-OCT-1999; | 990S-0161359. |
| PR | 26-OCT-1999; | 990S-0161360. |
| PR | 26-OCT-1999; | 990S-0161361. |
| PR | 28-OCT-1999; | 990S-0161920. |
| PR | 28-OCT-1999; | 990S-0161992. |
| PR | 29-OCT-1999; | 990S-0161993. |
| PR | 29-OCT-1999; | 990S-0162142. |

| | | | |
|--------------------------|--|--|-----|
| QY | 18 | FVNDCRQIGITCPIPVGIMPINNYKGFIRMTEGCKTKIIPADIMAALEPIDKDNEAVNAYG | 77 |
| Dd | 1 | FVNDRCRQIGINCPPIVPGLMPINNRYNGFLRMTEGCKTKIPVEYMALAEPIKDNEAVNAYG | 60 |
| QY | 78 | IHLGEMCKKLIAHGIKTLHLTYLLTMESALAIIMNLGLIEESKYSRLPWRPRANFRV | 137 |
| Dd | 61 | IHLGEMCKKLIAHGKLSLHYLTLMESALAIIMNLGLIDESKISRLPWRPRANFRV | 120 |
| QY | 138 | KEDVPPIFMANPKSYISTRTGMDOYPHGRMGDSGNPSYGALSDYOFMRPARDKLYEE | 197 |
| Dd | 121 | KEDVPPIFMANPKSYISTRTKGHEFPGRGRMGDSASASYGALSHDHSRRPARDKLOOE | 180 |
| QY | 198 | MAVPLKSVEDIYERFRLYLCLGRSNPMSSELDGIOPETKIINEOLEKINTKGFLTINSOP | 257 |
| Dd | 181 | MVWPLKSVEDIOEKFEKELCLGLNKSSPMSELDDGIOPETKIINEOLYVNKSGFLTINSQP | 240 |
| QY | 258 | AANGKSSSPYVGNGGGPGGYOYQKRAYVEFFCSKEKLDAVDKCDDRSLITYMANKGSW | 317 |
| Dd | 241 | SVNARSSDSPYVGWGGPVGYOYQKAYLEFFCSKEKLDAVBECKALPSITYMAVNGGEOW | 300 |
| QY | 318 | KSNVQTGVNNATVMGWVFPAKETIOPTIYDPVSFNWNKDEAFETISRGNASTYPPDEASRK | 377 |
| Dd | 301 | VSNVTQADVYNATVMGWVFPAKETIOPTIYDPSPFNWNKDEAFETISRGNASTYPPADESRN | 360 |
| QY | 378 | LVEEYGGSHFLVSLVDNDYINGDLFAVFAVD | 407 |
| Dd | 361 | LLEEYKNSTYSLVENDYINGDLFAVFAVD | 390 |
| RESULT 9 AAG15451 | | | |
| ID | AAG15451 standard; Protein; 373 AA. | | |
| XX | | | |
| AC | AAG15451; | | |
| XX | | | |
| DT | 17-OCT-2000 (first entry) | | |
| XX | | | |
| DE | Arabidopsis thaliana protein fragment SEQ ID NO: 15708. | | |
| XX | | | |
| KW | Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence. | | |
| KW | | | |
| XX | Arabidopsis thaliana. | | |
| OS | | | |
| PN | EP1033405-A2. | | |
| XX | | | |
| PD | 06-SEP-2000. | | |
| XX | | | |
| PF | 25-FEB-2000; 2000EP-0301439. | | |
| XX | | | |
| PR | 25-FEB-1999; | 99US-0121825. | |
| PR | 05-MAR-1999; | 99US-0123180. | |
| PR | 09-MAR-1999; | 99US-0123548. | |
| PR | 23-MAR-1999; | 99US-0125788. | |
| PR | 25-MAR-1999; | 99US-0126284. | |
| PR | 29-MAR-1999; | 99US-0126785. | |
| PR | 01-APR-1999; | 99US-0127462. | |
| PR | 06-APR-1999; | 99US-0128234. | |
| PR | 08-APR-1999; | 99US-0128714. | |
| PR | 16-APR-1999; | 99US-0129845. | |
| PR | 19-APR-1999; | 99US-0130077. | |
| PR | 21-APR-1999; | 99US-0130449. | |
| PR | 23-APR-1999; | 99US-0130810. | |
| PR | 23-APR-1999; | 99US-0130891. | |
| PR | 28-APR-1999; | 99US-0131449. | |
| PR | 30-APR-1999; | 99US-0132048. | |
| PR | 30-APR-1999; | 99US-0132407. | |
| PR | 04-MAY-1999; | 99US-0132484. | |
| PR | 05-MAY-1999; | 99US-0132485. | |
| PR | 06-MAY-1999; | 99US-0132486. | |
| PR | 06-MAY-1999; | 99US-0132487. | |

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.1%; Score 1664; DB 21; Length 373;

Best Local Similarity 81.5%; Pred. No. 3,6e-158;

Matches 303; Conservative 34; Mismatches 35; Indels 0; Gaps 0;

OY 36 MPNNKGFIRMTGFCFKTKIPADIMAALEPIKDNEBAVAYGIHGTETCKKTLAHGIKT 95
DB 1 MPNNKRGFLRMTEGFKTKIPYEVMAALEPIKDNEBAVAYGIHGTETCKKMLAHGVKS 60
OY 96 LHIYTLNMEKSALEIIMNGLIEESKVSRLPWRBPANVERKEDVPRPIFMANRKSYS 155
DB 61 LHIYTLNMEKSALEIIMNGLIEESKVSRLPWRBPANVERKEDVPRPIFMANRKSYS 120
OY 156 RTIGMDYPHGRMGDSNPSYALSDYOPMRPADKILVEEAVPLKSVEDIYERFRLY 215
DB 121 RTKGWEDFPOGRMGDSRSYALSDHOSRPRADKIQOEVMYPLKSVEDIQKFEEL 180
OY 216 CIGKILSNWSELDIQPTKILINOLEKINTKGFLTINSOPAVNGEKSDSPTVGWGGPG 275
DB 181 CIGKILSNWSELDIQPTKILINOLEKINTKGFLTINSOPAVNGEKSDSPTVGWGGPG 240
OY 276 GYVYOKAYVEFCSEKELDALVDCKDKDRTSLTYMAVNNKDSKSNVSGTDVNAVWGVFP 335
DB 241 GYVYOKAYVEFCSEKELDALVDCKDKDRTSLTYMAVNNKDSKSNVSGTDVNAVWGVFP 300
OY 336 AKETIIOPTVDPVSFNWVKDEAFELIWSRGWASLYPEDASRKLVEVGSHFLVSLVDND 395
DB 301 AKETIIOPTVDPVSFNWVKDEAFELIWSRGWASLYPEDASRKLVEVGSHFLVSLVDND 360
OY 396 YINGDLFAVFAAD 407
DB 361 YINGDLFAVFAAD 372

RESULT 10

AA615452

ID AA615452 standard; Protein: 362 AA.

AC AA615452;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 15709.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123160.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-01300449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135323.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

| | | |
|----|-------------|--------------|
| PR | 16-JUL-1999 | 9905-0144086 |
| PR | 15-JUL-1999 | 9905-0144325 |
| PR | 19-JUL-1999 | 9905-0144331 |
| PR | 19-JUL-1999 | 9905-0144332 |
| PR | 19-JUL-1999 | 9905-0144333 |
| PR | 19-JUL-1999 | 9905-0144334 |
| PR | 19-JUL-1999 | 9905-0144335 |
| PR | 20-JUL-1999 | 9905-0144352 |
| PR | 20-JUL-1999 | 9905-0144632 |
| PR | 20-JUL-1999 | 9905-0144864 |
| PR | 21-JUL-1999 | 9905-0144814 |
| PR | 21-JUL-1999 | 9905-0145086 |
| PR | 22-JUL-1999 | 9905-0145085 |
| PR | 22-JUL-1999 | 9905-0145087 |
| PR | 22-JUL-1999 | 9905-0145088 |
| PR | 22-JUL-1999 | 9905-0145192 |
| PR | 23-JUL-1999 | 9905-0145145 |
| PR | 23-JUL-1999 | 9905-0145148 |
| PR | 23-JUL-1999 | 9905-0145218 |
| PR | 23-JUL-1999 | 9905-0145224 |
| PR | 26-JUL-1999 | 9905-0145276 |
| PR | 27-JUL-1999 | 9905-0145913 |
| PR | 27-JUL-1999 | 9905-0145918 |
| PR | 27-JUL-1999 | 9905-0145919 |
| PR | 28-JUL-1999 | 9905-0145951 |
| PR | 02-AUG-1999 | 9905-0146386 |
| PR | 02-AUG-1999 | 9905-0146388 |
| PR | 02-AUG-1999 | 9905-0146389 |
| PR | 03-AUG-1999 | 9905-0147038 |
| PR | 04-AUG-1999 | 9905-0147204 |
| PR | 04-AUG-1999 | 9905-0147302 |
| PR | 05-AUG-1999 | 9905-0147192 |
| PR | 05-AUG-1999 | 9905-0147260 |
| PR | 06-AUG-1999 | 9905-0147303 |
| PR | 06-AUG-1999 | 9905-0147416 |
| PR | 09-AUG-1999 | 9905-0147493 |
| PR | 09-AUG-1999 | 9905-0147935 |
| PR | 10-AUG-1999 | 9905-0148171 |
| PR | 11-AUG-1999 | 9905-0148319 |
| PR | 12-AUG-1999 | 9905-0148341 |
| PR | 13-AUG-1999 | 9905-0148565 |
| PR | 13-AUG-1999 | 9905-0148684 |
| PR | 16-AUG-1999 | 9905-0149368 |
| PR | 17-AUG-1999 | 9905-0149172 |
| PR | 18-AUG-1999 | 9905-0149426 |
| PR | 20-AUG-1999 | 9905-0149722 |
| PR | 20-AUG-1999 | 9905-0149723 |
| PR | 23-AUG-1999 | 9905-0149902 |
| PR | 23-AUG-1999 | 9905-0149930 |
| PR | 23-AUG-1999 | 9905-0149936 |
| PR | 25-AUG-1999 | 9905-0150566 |
| PR | 26-AUG-1999 | 9905-0150884 |
| PR | 27-AUG-1999 | 9905-0151065 |
| PR | 27-AUG-1999 | 9905-0151066 |
| PR | 27-AUG-1999 | 9905-0151080 |
| PR | 30-AUG-1999 | 9905-0151303 |
| PR | 31-AUG-1999 | 9905-0151430 |
| PR | 01-SEP-1999 | 9905-0151930 |
| PR | 07-SEP-1999 | 9905-0152363 |
| PR | 10-SEP-1999 | 9905-0153070 |
| PR | 13-SEP-1999 | 9905-0153758 |
| PR | 15-SEP-1999 | 9905-0155656 |
| PR | 16-SEP-1999 | 9905-0154018 |
| PR | 16-SEP-1999 | 9905-0154039 |
| PR | 20-SEP-1999 | 9905-0154779 |
| PR | 22-SEP-1999 | 9905-0155133 |
| PR | 23-SEP-1999 | 9905-0155486 |
| PR | 24-SEP-1999 | 9905-0155659 |
| PR | 28-SEP-1999 | 9905-0156458 |
| PR | 29-SEP-1999 | 9905-0156596 |
| PR | 04-OCT-1999 | 9905-0157117 |
| PR | 05-OCT-1999 | 9905-0157553 |
| PR | 06-OCT-1999 | 9905-0157865 |
| PR | 07-OCT-1999 | 9905-0158029 |

| | | |
|----|--------------|---------------|
| PR | 08-OCT-1999; | 9905-0156233; |
| PR | 12-OCT-1999; | 9905-0156369; |
| PR | 13-OCT-1999; | 9905-0156293; |
| PR | 13-OCT-1999; | 9905-0155294; |
| PR | 13-OCT-1999; | 9905-0155295; |
| PR | 14-OCT-1999; | 9905-0155329; |
| PR | 14-OCT-1999; | 9905-0155330; |
| PR | 14-OCT-1999; | 9905-0155331; |
| PR | 14-OCT-1999; | 9905-0155637; |
| PR | 14-OCT-1999; | 9905-0155638; |
| PR | 18-OCT-1999; | 9905-0155584; |
| PR | 21-OCT-1999; | 9905-0165741; |
| PR | 21-OCT-1999; | 9905-0160767; |
| PR | 21-OCT-1999; | 9905-0160768; |
| PR | 21-OCT-1999; | 9905-0160770; |
| PR | 21-OCT-1999; | 9905-0160815; |
| PR | 21-OCT-1999; | 9905-0160815; |
| PR | 22-OCT-1999; | 9905-0160980; |
| PR | 22-OCT-1999; | 9905-0160981; |
| PR | 22-OCT-1999; | 9905-0160989; |
| PR | 22-OCT-1999; | 9905-0161404; |
| PR | 25-OCT-1999; | 9905-0161405; |
| PR | 25-OCT-1999; | 9905-0161406; |
| PR | 26-OCT-1999; | 9905-0161355; |
| PR | 26-OCT-1999; | 9905-0161360; |
| PR | 26-OCT-1999; | 9905-0161361; |
| PR | 28-OCT-1999; | 9905-0161920; |
| PR | 28-OCT-1999; | 9905-0161993; |
| PR | 29-OCT-1999; | 9905-0162142; |

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 73.5%; | Score 1608; | DB 21; | Length 362; |
| Best Local Similarity | 81.4%; | Pred. No. 1.4e-152; | | |
| Matches 294; | Conservative 35; | Mismatches 35; | Indels 0; | Gaps 0; |

[illegible]

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126284.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127452.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-01300449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137232.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139432.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.

PR 23-JUN-1999; 990S-0140353.
PR 24-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140693.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142380.
PR 09-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145291.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.

[illegible]

| | |
|-----------|---|
| RESULT 12 | |
| AAAG36170 | |
| ID | AAAG36170 standard; Protein; 470 AA. |
| XX | |
| AC | AAAG36170; |
| XX | |
| DT | 18-OCT-2000 (first entry) |
| XX | |
| DE | Arabidopsis thaliana protein fragment SEQ ID NO: 44289. |
| XX | |
| KM | Protein identification; signal transduction pathway; metabolic pathway; |
| KM | hybridisation assay; genetic mapping; gene expression control; promoter |
| KM | termination sequence. |
| XX | |
| OS | Arabidopsis thaliana. |
| XX | |
| PN | EP1033405-A2. |
| XX | |
| PD | 06-SEP-2000. |
| XX | |
| PF | 25-FEB-2000; 2000EP-0301439. |
| XX | |
| PR | 25-FEB-1999; 99US-0121825. |
| PR | 05-MAR-1999; 99US-0123180. |
| PR | 09-MAR-1999; 99US-0123548. |
| PR | 23-MAR-1999; 99US-0125788. |
| PR | 25-MAR-1999; 99US-0126264. |
| PR | 29-MAR-1999; 99US-0126785. |
| PR | 01-APR-1999; 99US-0127462. |
| PR | 06-APR-1999; 99US-0128234. |
| PR | 08-APR-1999; 99US-0128714. |
| PR | 16-APR-1999; 99US-0129845. |
| PR | 19-APR-1999; 99US-0130077. |
| PR | 21-APR-1999; 99US-0130449. |
| PR | 23-APR-1999; 99US-0130510. |
| PR | 23-APR-1999; 99US-0130891. |
| PR | 28-APR-1999; 99US-0131449. |
| PR | 30-APR-1999; 99US-0132048. |
| PR | 30-APR-1999; 99US-0132407. |
| PR | 04-MAY-1999; 99US-0132484. |
| PR | 05-MAY-1999; 99US-0132485. |
| PR | 06-MAY-1999; 99US-0132486. |
| PR | 06-MAY-1999; 99US-0132487. |
| PR | 07-MAY-1999; 99US-0132863. |
| PR | 11-MAY-1999; 99US-0134256. |
| PR | 14-MAY-1999; 99US-0134218. |
| PR | 14-MAY-1999; 99US-0134219. |
| PR | 14-MAY-1999; 99US-0134221. |
| PR | 14-MAY-1999; 99US-0134370. |
| PR | 18-MAY-1999; 99US-0134768. |
| PR | 19-MAY-1999; 99US-0134941. |
| PR | 20-MAY-1999; 99US-0135124. |
| PR | 21-MAY-1999; 99US-0135353. |
| PR | 24-MAY-1999; 99US-0135629. |
| PR | 25-MAY-1999; 99US-0136021. |
| PR | 27-MAY-1999; 99US-0136392. |
| PR | 28-MAY-1999; 99US-0136782. |
| PR | 03-JUN-1999; 99US-0137222. |
| PR | 03-JUN-1999; 99US-0137528. |
| PR | 04-JUN-1999; 99US-0137502. |
| PR | 07-JUN-1999; 99US-0137724. |
| PR | 08-JUN-1999; 99US-0138094. |
| PR | 10-JUN-1999; 99US-0138540. |
| PR | 10-JUN-1999; 99US-0138847. |
| PR | 14-JUN-1999; 99US-0139119. |
| PR | 16-JUN-1999; 99US-0139452. |
| PR | 17-JUN-1999; 99US-0139492. |
| PR | 18-JUN-1999; 99US-0139454. |
| PR | 18-JUN-1999; 99US-0139455. |
| PR | 18-JUN-1999; 99US-0139456. |
| PR | 18-JUN-1999; 99US-0139457. |
| PR | 18-JUN-1999; 99US-0139458. |

[illegible]

| | | |
|----|-------------|--------------|
| PR | 16-JUN-1999 | 9905-0139452 |
| PR | 16-JUN-1999 | 9905-0139453 |
| PR | 17-JUN-1999 | 9905-0139454 |
| PR | 18-JUN-1999 | 9905-0139455 |
| PR | 18-JUN-1999 | 9905-0139456 |
| PR | 18-JUN-1999 | 9905-0139457 |
| PR | 18-JUN-1999 | 9905-0139458 |
| PR | 18-JUN-1999 | 9905-0139459 |
| PR | 21-JUN-1999 | 9905-0139817 |
| PR | 22-JUN-1999 | 9905-0139859 |
| PR | 23-JUN-1999 | 9905-0140353 |
| PR | 24-JUN-1999 | 9905-0140354 |
| PR | 25-JUN-1999 | 9905-0140695 |
| PR | 06-JUL-1999 | 9905-0142390 |
| PR | 08-JUL-1999 | 9905-0142803 |
| PR | 09-JUL-1999 | 9905-0142920 |
| PR | 12-JUL-1999 | 9905-0142977 |
| PR | 13-JUL-1999 | 9905-0143547 |
| PR | 14-JUL-1999 | 9905-0143624 |
| PR | 15-JUL-1999 | 9905-0144005 |
| PR | 16-JUL-1999 | 9905-0144085 |
| PR | 16-JUL-1999 | 9905-0144086 |
| PR | 19-JUL-1999 | 9905-0144325 |
| PR | 19-JUL-1999 | 9905-0144331 |
| PR | 19-JUL-1999 | 9905-0144332 |
| PR | 19-JUL-1999 | 9905-0144333 |
| PR | 19-JUL-1999 | 9905-0144334 |
| PR | 19-JUL-1999 | 9905-0144335 |
| PR | 20-JUL-1999 | 9905-0144332 |
| PR | 20-JUL-1999 | 9905-0144632 |
| PR | 20-JUL-1999 | 9905-0144684 |
| PR | 21-JUL-1999 | 9905-0144814 |
| PR | 21-JUL-1999 | 9905-0145086 |
| PR | 21-JUL-1999 | 9905-0145088 |
| PR | 21-JUL-1999 | 9905-0145085 |
| PR | 22-JUL-1999 | 9905-0145087 |
| PR | 22-JUL-1999 | 9905-0145089 |
| PR | 22-JUL-1999 | 9905-0145112 |
| PR | 23-JUL-1999 | 9905-0145218 |
| PR | 23-JUL-1999 | 9905-0145224 |
| PR | 26-JUL-1999 | 9905-0145276 |
| PR | 27-JUL-1999 | 9905-0145913 |
| PR | 27-JUL-1999 | 9905-0145918 |
| PR | 27-JUL-1999 | 9905-0145919 |
| PR | 28-JUL-1999 | 9905-0145951 |
| PR | 28-AUG-1999 | 9905-0146386 |
| PR | 02-AUG-1999 | 9905-0146388 |
| PR | 02-AUG-1999 | 9905-0146389 |
| PR | 03-AUG-1999 | 9905-0147034 |
| PR | 04-AUG-1999 | 9905-0147204 |
| PR | 04-AUG-1999 | 9905-0147302 |
| PR | 05-AUG-1999 | 9905-0147192 |
| PR | 05-AUG-1999 | 9905-0147260 |
| PR | 06-AUG-1999 | 9905-0147403 |
| PR | 06-AUG-1999 | 9905-0147416 |
| PR | 09-AUG-1999 | 9905-0147493 |
| PR | 09-AUG-1999 | 9905-0147935 |
| PR | 10-AUG-1999 | 9905-0148171 |

PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148365.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 25-AUG-1999; 9905-0149930.
PR 26-AUG-1999; 9905-0150566.
PR 27-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 30-AUG-1999; 9905-0151080.
PR 31-AUG-1999; 9905-0151303.
PR 01-SEP-1999; 9905-0151338.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

| | | | | | |
|----|---------|--|-------------|---------------------|-----------------------|
| | | 54.1%; | Score 1184; | DB 21, | Length 497; |
| | | Best Local Similarity | 78.9%; | Pred. NO. 9.7e-110; | |
| | Matches | 213; Conservative | 26; | Mismatches | 29; Indels 0; Gaps 0; |
| Oy | 1 | AATVAVQLPPYDDIDIFKTPVNCRCQIGTCPLPVGIMPIPNKYGRIRMTGFCCKRIAPADIM | 60 | | |
| | | : : : : : : | : | | |
| Dd | 214 | ADDIVQLERFDVIDFKFVNDRQGICNPFLPGIIMPISNNKGILRNAAGCCKRIPAEALT | 273 | | |
| | | : : : : : : | : | | |

| | | | |
|----|-----|--|-----|
| 0Y | 61 | ALAEIKONEEVKXVYGHLETCMOCKILLANGITLHLYTLNMRKSALATIMNGLIEES | 120 |
| Db | 274 | AALEBIKONDEWVYGVHFEATCKCKILLANGITSLHLYTLNVRKSXIGIMNGLIDES | 333 |
| 0Y | 121 | KVSRSLPMRRPANYRVKEDVRPIFWANRPKSYISRTIGMDQYHGMCSOCPNSVALS | 160 |
| Db | 334 | KISRSLPMRRPANYRTEVEDVRPIFWANRPKSYISRTIKGMNDPFHGMGDSHSAVSTLS | 393 |
| 0Y | 161 | DYOFMRPBARDDKILVEEMAVPLKSVEDYIERFRLYCLGKLRNSPWSLSDLOPETKLINE | 240 |
| Db | 394 | DYOFMRPBARDDKILQOEEMVPLKSIEDYQEKFKELGJNLKSPWSLSDLOPETKLINE | 453 |
| 0Y | 241 | OLEKINTKGFLINSQPAVNGEKSDSPVVG | 270 |
| Db | 454 | OLGKINSNGFLINSQPSVNAKSDSPAIG | 483 |

```

RESULT 14
AAR8358
ID AAR8358 standard; Protein; 656 AA.
XX
AC
XX AAR8358;
XX
DT 15-OCT-1995 (first entry)
XX
DE Human methylene-tetrahydrofolate-reductase.
XX
KW Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;
RV cardiovascular disease; neurological disease; folic acid metabolism;
KW EC:1.5.1.20; enzyme.
XX
OS Homo sapiens.
XX
PN W09533054-A1.
XX
PD 07-DEC-1995.
XX
PF 25-MAY-1995; 95WO-CA00314.
XX
PR 26-MAY-1994; 94GB-0010620.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Goyette P, Rozen R;
XX
DR MPI; 1996-030565/O3.
XX
DR N-PDB; AAT09689.
XX
PT Human methylene:tetra:hydro:folate reductase cDNA probe - for
PT detection of sequence abnormalities in methylene:tetra:hydro:folate
PT reductase e.g. in cardiovascular, neurological or folic acid
PT metabolism disorders
PS
XX Disclosure: Fig.6A-6C; 66pp; English.
CC This is the protein sequence encoding human MTHFR, the gene of
CC which has been localised to chromosome 1p36.3. Deficiencies of
CC this protein may lead to cardiovascular and neurological disorders
CC and disorders influences by folic acid metabolism.
XX
SQ Sequence 656 AA.
Query Match 44.8%; Score 979; DB 17; Length 656;
Best Local Similarity 44.9%; Pred. No. 6e-89;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps
OY 1 ARVAVQLGVDDDIPLAKFNOCROGLTCIPYPGIMPLNMGKTRTGTGCKTKRPADIM 60
| :||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 222 ADRIITQLEFADFRRFYKACTDGMGITCPVPGLPFGIYGHSIROVKLSLDEVQEIK 281
| ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
:|||||KDNEEAVKAYGIHLGTEMCKRTILAG-IKTLLHLYLNMEKSALAILNMNGIEE 119

```

```

Db      262 DVIEPIKNDAAIRNIGIELAVSLCOELLASGLVPGLFHYTLINREMATTEVLRLGMWTE 341
Qy      120 SKVSRSLPWRPARRANFRVKEDVRPIFWANRPKSYISRTIGMDOYPHGRMGDSCNPSTGAL 179
        ||| | : : : ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      342 DP-RRLPFWALSAHPKRREEDVRPIFWASRPKSYIYRTOEWDEFPNGRNGSSSPAFGEL 400
Qy      180 SDYQ--FMRRPARDKKLYEEMAVPLKSYEDIYERFLXCGKLRSN-----PMSELDG 230
        || : : : : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      401 KDYLFLYLSKSPKPEELLKMGEBELTSEASVEFVLYLSEPNRNGHKVYCLPAMD-EP 459
Qy      231 LOPETKIINEOLEKINTKGFLTINSOPAVNGEKSDSPYVGWGGPGGYVYOKAYVEFFCSK 290
        | : : : : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      460 LAAETSLKEELLRVNRGILITINSQPNINGKPSDPIVGMGSPSGGYVFOKAYLEFFTSR 519
Qy      291 EKDALVDKCD-RTSLTYMAVANKGSKSNVGGTQDVNAVTVGVFPAKEIIQPTIVDPVS 349
        | : : : : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      520 ETAEALLOVLYKKEELRVNHLVNVKSGENTNAPLOPNNAVTVGIFPGREIIQPTIVDPVS 579
Qy      350 FNMWKEAFELMSRGNWASLYPEDEASRKLYEEVGGSHFLVSLVNDY-INGDLFAVFPAD 407
        | : : : : : | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      580 FNMWKEAFALMTERMGKLYEESPSRTIIQYIHDNIFLVNLVNDPPLDNCIMQVVED 638

RESULT 15
AAE12607
ID      AAE12607 standard; Protein; 656 AA.
XX
XX      AAE12607;
XX
Df      03-JAN-2002 (first entry)
XX
XX      Human methylenetetrahydrofolate reductase (MTHFR) protein #2.
XX
XX      Human: methylenetetrahydrofolate reductase; MTHFR; gene therapy; glioma;
XX      anticense therapy; EC 1.5.1.20; chromosome 1p36.3; cancer; kidney cell;
XX      pancreas; lung; brain; cytosolic; colon carcinoma; breast; colorectal;
XX      neuroblastoma; leukaemia.
XX
OS      Homo sapiens.
XX
XX      US2001025030-A1.
XX
XX      27-SEP-2001.
XX
XX      01-DEC-2000; 2000US-0728910.
XX
XX      01-MAR-1999; 99US-0258928.
XX
XX      (ROZE/) ROZEN R.
XX      PA
XX      (SEKH/) SEKHON J.
XX
XX      Rozen R, Sekhon J;
XX
XX      WPI; 2001-638509/73.
XX      DR
XX      N-PSDB; AAD20463.
XX
XX      New antisense nucleic acids, which are methylenetetrahydrofolate
XX      PT reductase inhibitors, useful for treating, stabilizing or preventing
XX      PT cancer, e.g. breast carcinoma, colon carcinoma, colorectal carcinoma or
XX      neuroblastoma
XX
XX      Disclosure; Fig 6; 68pp; English.
XX
XX      The invention relates to a non allele-specific antisense nucleic acids,
XX      CC which binds to methylenetetrahydrofolate reductase (MTHFR; EC 1.5.1.20)
XX      CC nucleic acids and inhibits MTHFR protein expression in a mammal. MTHFR
XX      CC catalyses the NADPH-linked reduction of 5,10-methylenetetrahydrofolate to
XX      CC 5-methylenetetrahydrofolate, a co-substrate for methylation of
XX      CC homocysteine to methionine. The invention provides potential therapy for
XX      CC individuals with MTHFR deficiency. The non allele-specific antisense
XX      CC nucleic acids are useful for treating, stabilizing or preventing cancer,
XX      CC particularly breast carcinoma, colon carcinoma, colorectal carcinoma,
XX      CC lung cancer, brain cancer, pancreatic cancer, kidney cell cancer.

```

```

CC      neuroblastoma, glioma and leukaemia. MTHFR cDNA probe is used in gene
CC      therapy. The present sequence is human methylenetetrahydrofolate
CC      reductase (MTHFR) protein #2. Human MTHFR gene is mapped to
CC      chromosome 1p36.3.
XX
XX      Sequence 656 AA;
SQ
Query Match 44.8%; Score 979; DB 22; Length 656;
Best Local Similarity 44.9%; Pred. No. 6e-89;
Matches 188; Conservative 79; Mismatches 138; Indels 14; Gaps 7;
Qy      1 ARVITQLEFVDTDFLFKFNVDCKRQIGTCPIVRSIMINNYKGFIRMTGCKTKIPADIM 60
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      222 ADPIITQLEFFADDFPRVAVACTDMGITCPVPGIFIQHSLSRLQVLSKLEVPQETK 281
Qy      61 AALEPIKNEAVKAYGIGHGTEMCKKILAHG-IKTLHLTYLNNKESALAILNMLGIEE 119
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      282 DVIEPIKNDAAIRNIGIELAVSLCOELLASGLVPGLFHYTLINREMATTEVLRLGMWTE 341
Qy      120 SKVSRSLPWRPARRANFRVKEDVRPIFWANRPKSYISRTIGMDOYPHGRMGDSCNPSTGAL 179
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      342 DP-RRLPFWALSAHPKRREEDVRPIFWASRPKSYIYRTOEWDEFPNGRNGSSSPAFGEL 400
Qy      180 SDYQ--FMRRPARDKKLYEEMAVPLKSYEDIYERFLXCGKLRSN-----PMSELDG 230
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      401 KDYLFLYLSKSPKPEELLKMGEBELTSEASVEFVLYLSEPNRNGHKVYCLPAMD-EP 459
Qy      231 LOPETKIINEOLEKINTKGFLTINSOPAVNGEKSDSPYVGWGGPGGYVYOKAYVEFFCSK 290
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      460 LAAETSLKEELLRVNRGILITINSQPNINGKPSDPIVGMGSPSGGYVFOKAYLEFFTSR 519
Qy      291 EKDALVDKCD-RTSLTYMAVANKGSKSNVGGTQDVNAVTVGVFPAKEIIQPTIVDPVS 349
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      520 ETAEALLOVLYKKEELRVNHLVNVKSGENTNAPLOPNNAVTVGIFPGREIIQPTIVDPVS 579
Qy      350 FNMWKEAFELMSRGNWASLYPEDEASRKLYEEVGGSHFLVSLVNDY-INGDLFAVFPAD 407
        | : : : : : | : : | : : | | | | | | | | | | | | | | | | | | | | | | |
Db      580 FNMWKEAFALMTERMGKLYEESPSRTIIQYIHDNIFLVNLVNDPPLDNCIMQVVED 638

```

Search completed: February 14, 2003, 23:07:42
Job time : 73 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 14, 2003, 19:26:12 : Search time 22 Seconds

(without alignments)
769.197 Million cell updates/sec

Title: US-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVYVQLFDTIDFLKFN.....VSLVDNDYINGDLFAVADF 408

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 1834 | 83.9 | 594 | 1 MTHR_ARATH | O08585 arabidopsis |
| 2 | 979 | 44.8 | 656 | 1 MTHR_HUMAN | P42898 homo sapien |
| 3 | 943 | 43.1 | 654 | 1 MTHR_MOUSE | O9WU20 mus musculu |
| 4 | 895 | 40.9 | 603 | 1 MTHR_SCHPO | Q10258 schizosacch |
| 5 | 887 | 40.6 | 663 | 1 MTHR_CAEEL | Q17693 caenorhabdi |
| 6 | 778 | 35.6 | 599 | 1 MTHS_YEAST | P53128 saccharomyc |
| 7 | 629 | 28.8 | 641 | 1 MTHS_SCHPO | O79927 schizosacch |
| 8 | 578.5 | 26.5 | 657 | 1 MTHR_YEAST | P46151 saccharomyc |
| 9 | 222 | 10.2 | 296 | 1 MTHR_YEAST | P46151 saccharomyc |
| 10 | 193 | 8.8 | 307 | 1 MTHR_STRLI | O54232 aquilex aeo |
| 11 | 189 | 8.6 | 292 | 1 MTHR_HAEIN | P45208 haemophilus |
| 12 | 189 | 8.6 | 296 | 1 MTHR_ECOLI | P00394 escherichia |
| 13 | 184 | 8.4 | 296 | 1 MTHR_SALTY | P11003 salmonella |
| 14 | 176 | 8.0 | 298 | 1 MTHR_ERWCA | P71319 erwina car |
| 15 | 168 | 7.7 | 292 | 1 MTHR_BUCAL | P57154 buchiera ap |
| 16 | 105.5 | 4.8 | 2156 | 1 RPL_HUMAN | P56715 homo sapien |
| 17 | 105 | 4.6 | 684 | 1 YMW5_YEAST | O04779 saccharomyc |
| 18 | 101 | 4.6 | 548 | 1 PHRL_CANAL | P43076 candida alb |
| 19 | 99 | 4.5 | 422 | 1 NCAP_VSVIG | P11212 vesicular s |
| 20 | 98.5 | 4.5 | 1067 | 1 CARB_CLOPE | O84HB3 clostridium |
| 21 | 97 | 4.4 | 422 | 1 NCAP_VSVST | P03521 vesicular s |
| 22 | 96 | 4.4 | 900 | 1 SYA_MYCPN | P75368 mycoplasma |
| 23 | 96 | 4.4 | 924 | 1 CSKP_MOUSE | O70589 mus musculu |
| 24 | 95 | 4.3 | 1005 | 1 VIA_PSVJ | P28726 peanut stru |
| 25 | 94 | 4.3 | 921 | 1 CSKP_HUMAN | O19936 homo sapien |
| 26 | 93.5 | 4.3 | 437 | 1 KDTA_CHLHN | Q46222 chlamydia p |
| 27 | 92.5 | 4.2 | 1231 | 1 BLM_CAEEL | O18017 caenorhabdi |
| 28 | 92 | 4.2 | 879 | 1 PHSG_AOUAE | O66932 aquilex aeo |
| 29 | 91.5 | 4.2 | 879 | 1 MAMB_CAPII | O95327 capri hircu |
| 30 | 91 | 4.2 | 886 | 1 DSCI_MOUSE | P55849 mus musculu |
| 31 | 91 | 4.2 | 1026 | 1 MYIB_DROME | O29979 drosophila |
| 32 | 91 | 4.2 | 1161 | 1 DP3A_AOUAE | O61125 aquilex aeo |
| 33 | 90.5 | 4.1 | 434 | 1 PNCB_RHIO | O98d24 rhizobium 1 |

| | | | | | |
|----|------|-----|------|--------------|--------------------|
| 34 | 90.5 | 4.1 | 568 | 1 PPO1_AGABI | O00024 agaricus bi |
| 35 | 90 | 4.1 | 391 | 1 POLG_CYVV | P20177 clover yell |
| 36 | 90 | 4.1 | 549 | 1 EPD2_CANMA | O74137 candida mal |
| 37 | 89.5 | 4.1 | 387 | 1 RH18_SCHPO | O74747 schizosacch |
| 38 | 89.5 | 4.1 | 639 | 1 CALC_RABIT | Q28902 oryctolagus |
| 39 | 89.5 | 4.1 | 3119 | 1 CALC_MOUSE | Q06847 mus musculu |
| 40 | 89 | 4.1 | 446 | 1 Y612_METJA | O58029 methanococ |
| 41 | 89 | 4.1 | 494 | 1 SYK_BACST | O9rhv9 bacillus st |
| 42 | 89 | 4.1 | 610 | 1 LKHA_MOUSE | P24527 mus musculu |
| 43 | 88.5 | 4.0 | 572 | 1 SYM_AERPE | O9cy3 aeropyrum p |
| 44 | 88.5 | 4.0 | 1466 | 1 SPA2_YEAST | P23201 saccharomyc |
| 45 | 88 | 4.0 | 909 | 1 CSKP_RAT | O62915 rattus norv |

ALIGNMENTS

RESULT 1
MTHR_ARATH STANDARD: PRT: 594 AA.
AC O08585:09SEP59:094JZ1:09SU0;
DT 15-JUN-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylenetetrahydrofolate reductase (EC 1.5.1.20) (MTHFR2).
GN MTHFR1 OR ATG44160 OR F6E13.29.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae;
OC eucosids II: Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Ravelle S., Rebelle F., Douce R.;
RT "Folate metabolism in higher plants: cloning of a cDNA for
RT 5,10-methylenetetrahydrofolate reductase in Arabidopsis thaliana.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA MEDLINE=20062814; PubMed=10593891;
RA Roye S., Wang H., McNeil S.D., Raymond R.K., Appling D.R.,
RA Shachar-Hill Y., Boenert H.J., Hanson A.D.;
RT "Isolation, characterization, and functional expression of cDNAs
RT encoding NADH-dependent methylenetetrahydrofolate reductase from
RT higher plants.";
RL J. Biol. Chem. 274:36089-36096(1999).
[3]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldguy T.V.,
RA Buell K.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Motil K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
RA Tallon L.J., Gail J.E., Adams M.D., Carreira A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
[4]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC methylenetetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (By Similarity).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE

```

CC      (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ245414; CAB53783.1; -
DR      EMBL; AF181967; AAD55788.1; -
DR      EMBL; AC004005; AAC23420.2; -
DR      EMBL; AY050434; AAK91450.1; -
DR      EMBL; AF370515; AAK43892.1; -
DR      HSSP; P00394; 1B5T.
DR      InterPro; IPR004621; Fadh2_euk.
DR      InterPro; IPR003171; Methylcrof_reductase.
DR      Pfam; PF02219; MTHR.1.
DR      TIGRFAMs; TIGR00677; fadh2_euk; 1.
KW      Oxidoreductase; Flavoprotein; FAD; NADP.
FT      CONFLICT 72 N -> S (IN REF. 1).
FT      CONFLICT 115 P -> H (IN REF. 4; AAK3892).
FT      CONFLICT 508 A -> V (IN REF. 1).
SQ      SEQUENCE 594 AA; 66802 MW; E2F4F919FCE10EF8 CRC64;

Query Match      83.9% Score 1834; DB 1; Length 594;
Best Local Similarity 82.1%; Pred. No. 1,2e-141;
Matches 334; Conservative 37; Mismatches 36; Indels 0; Gaps 0;

QY      1 ARVITQLFYPDIDFLKFNDCRQIGTCPIPGIMPINNYKGFIRMTGFCCKIPADIM 60
DB      187 ADLIVTQLFYPDIDFLKFNDCRQIGTSCPIPGIMPINNYKGFIRMTGFCCKIPAEVM 246

QY      61 AALPEIKNEAAYAGYGHGTECKKLHAGITLHLYTNMKSALAIIMNLIEES 120
DB      247 AALPEIKNEAAYAGYGHGTECKKLHAGITLHLYTNMKSALAIIMNLGIDES 306

QY      121 KYSRSLPMPRPANVRYVEDRPIFMANRPKSYTSRTIGMDQYHGRMGDCSNPYGALS 180
DB      307 KISRSLPMPRPANVRYVEDRPIFMANRPKSYTSRTIGMEDPFGKMGDSRSAYGALS 366

QY      181 DYQMRPPARPKLVEEMAVPLKSVEDIYERFLYCLGKLSNPWSELGLQPTKTIINE 240
DB      367 DHQMRPPARPKLQEQEAVVPLKSVEDIQEKFKELGKLSNPWSELGLQPTKTIINE 426

QY      241 QLEKNTGELTINSQPAVNGEKSDPTVWGCGGYGAYVEFFCSREKDALVDKC 300
DB      427 QLIKNSGELTINSQPAVNGEKSDPTVWGCGGYGAYVEFFCSREKDALVDK 486

QY      301 KDRTSLTYMAVANKDGMKSNVGTDVNAVTVGVPFAKEITIOPTIVDPVSFWMKDEAFEI 360
DB      487 KALPSITYMAVANKGEQWNTAQADVNAVTVGVPFAKEITIOPTIVDPVSFWMKDEAFET 546

QY      361 WSRGMASLTPPEDEASRKLVEEVGGSHPLVSLVNDYINGDLFAVFAAD 407
DB      547 WSRGMANLYPEADPSRNILVEEKSSYLVSLVENDYINGDLFAVFAAD 593

RESULT 2
MTHR_HUMAN STANDARD; PRT; 656 AA.
AC P42898; Q9UOR2;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylcrofetrofolate reductase (EC 1.5.1.20).
GN MTHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP      SEQUENCE OF 1-415 FROM N.A., AND VARIANT GLN-157.
RC TRISUE-Liver;
RX MEDLINE=95004587; PubMed=7920641;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Matthews R.G., Rozen R.;
RT "Human methylcrofetrofolate reductase: isolation of cDNA,
RL mapping and mutation identification.";
RL Nat. Genet. 7:195-200(1994).
RN [12]
RP ERRATUM.
RX MEDLINE=95038846; PubMed=7951330;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Matthews R.G., Rozen R.;
RL Nat. Genet. 7:551-551(1994).
RN [13]
RP SEQUENCE FROM N.A.
RA Rozen R., Goyette P.;
RT "cDNA for human methylcrofetrofolate reductase.";
RL Patent number WO9533054, 07-DEC-1995.
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=98345426; PubMed=9680386;
RA Goyette P., Pal A., Milos R., Frosst P., Trian P., Chen Z., Chan M.,
RA Rozen R.;
RT "Gene structure of human and mouse methylcrofetrofolate reductase
RT (MTHR).";
RL Mamm. Genome 9:652-656(1998).
RN [15]
RP SEQUENCE FROM N.A.
RA Homberger A., Lindebank M., Winter C., Rapp B., Koch H.G.;
RT "Revised translation initiation site of the human
RL methylcrofetrofolate reductase (MTHR).";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [16]
RP VARIANTS Q-52; M-227; L-251; C-325; C-335 AND C-357, AND REVISION TO
RP 177.
RX MEDLINE=95243219; PubMed=7726158;
RA Goyette P., Frosst P., Rosenblatt D.S., Rozen R.;
RT "Seven novel mutations in the methylcrofetrofolate reductase
RT gene and genotype/phenotype correlations in severe
RL methylcrofetrofolate reductase deficiency.";
RL Am. J. Hum. Genet. 56:1052-1059(1995).
RN [17]
RP VARIANT VAL-222.
RX MEDLINE=95375773; PubMed=7647779;
RA Frosst P., Blom H.J., Milos R., Goyette P., Sheppard C.A.,
RA Matthews R.G., Boers G.J.H., den Heijer M., Kluitmans L.A.J.,
RA van den Heuvel L.P., Rozen R.;
RT "A candidate genetic risk factor for vascular disease: a common
RL mutation in methylcrofetrofolate reductase.";
RL Nat. Genet. 10:111-113(1995).
RN [18]
RP VARIANTS PRO-51; PRO-323 AND CYS-377.
RX MEDLINE=97094184; PubMed=8940272;
RA Goyette P., Christensen B., Rosenblatt D.S., Rozen R.;
RT "Severe and mild mutations in cis for the methylcrofetrofolate
RT reductase (MTHR) gene, and description of five novel mutations in
RT MTHR.";
RL Am. J. Hum. Genet. 59:1268-1275(1996).
RN [19]
RP VARIANT VAL-222.
RX MEDLINE=98213590; PubMed=9545406;
RA Schneider J.A., Rees D.C., Liu Y.-T., Clegg J.B.;
RT "Worldwide distribution of a common methylcrofetrofolate
RT reductase mutation.";
RL Am. J. Hum. Genet. 62:1258-1260(1998).
RN [110]
RP VARIANT ALA-428.
RX MEDLINE=98213566; PubMed=9545395;
RA van der Put N.M.J., Gabreels F., Stevens E.M.B., Smeltink J.A.M.,
RA Trijbels F.J.M., Eskes T.K.A.B., van den Heuvel L.P., Blom H.J.;
RT "A second mutation in the methylcrofetrofolate reductase gene: an
RT additional risk factor for neural-tube defects?";

```


Db 580 FFWKDEAFALMIERMGKLYEESPSRTIIQYIHNDNYFLVNLVNDPELNDCLMQVED 638

RESULT 3

MTNR_MOUSE STANDARD; PRT; 654 AA.

ID MTNR_MOUSE

AC 09JUN20;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Methylene tetrahydrofolate reductase (EC 1.5.1.20).

GN MTNR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98345426; PubMed=9680386;

RA Goyette P., Pal A., Milos R., Frosst P., Tran P., Chen Z., Chan M., Rozen R.;

RT "Gene structure of human and mouse methylenetetrahydrofolate reductase (MTHFR).";

RL Mamm. Genome 9:652-656(1998).

CC -1- FUNCTION: CATALYZES THE CONVERSION OF 5,10-METHYLENETHETRAHYDROFOLATE TO 5-METHYLTETRAHYDROFOLATE, A CO-SUBSTRATE FOR HOMOCYSTEINE REMETHYLATION TO METHIONINE.

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-methylenetetrahydrofolate + NADPH.

CC -1- COFACTOR: FAD (BY SIMILARITY).

CC -1- ENZYME REGULATION: ALLOSTERICALLY REGULATED BY S-ADENOSYLMETHIONINE (BY SIMILARITY).

CC -1- PATHWAY: Folate metabolism.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE METHYLENETHETRAHYDROFOLATE REDUCTASE (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF105988; AAD20313.1; JOINED.

DR EMBL: AF105989; AAD20313.1; JOINED.

DR EMBL: AF105990; AAD20313.1; JOINED.

DR EMBL: AF105991; AAD20313.1; JOINED.

DR EMBL: AF105992; AAD20313.1; JOINED.

DR EMBL: AF105993; AAD20313.1; JOINED.

DR EMBL: AF105994; AAD20313.1; JOINED.

DR EMBL: AF105995; AAD20313.1; JOINED.

DR EMBL: AF105996; AAD20313.1; JOINED.

DR EMBL: AF105997; AAD20313.1; JOINED.

DR HSSP: P00394; 1B5T.

DR MGP: MGI106639; Mthfr.

DR InterPro: IPR004621; Fadh2_euk.

DR InterPro: IPR003171; Mchdrof_redctse.

DR Pfam: PF02219; MTHFR_1.

DR TIGRFAMs: TIGR00677; fadh2_euk; 1.

KM Oxidoreductase; Flavoprotein; FAD; NADP.

SQ SEQUENCE 654 AA; 74649 MW; 12AD31806B371E17 CRC64;

Query Match 43.1%; Score 943; DB 1; Length 654;

Best Local Similarity 43.0%; Pired. No. 4,5e-69;

Matches 180; Conservative 84; Mismatches 141; Indels 14; Gaps 7;

QY 1 ARVLTQLFDTDFLEKFNVDRCQIGITCPVPGIMINNYKFFIMTGCKTKPIADIM 60

Db 221 ADFIITQLFDEASTFFSFVAKCEIGISCIPLGIFPIQGYTSLRLQVLNLSKLEVFQKIK 280

QY 61 AALEPIKDNKEAVKAYGHIHGTEMCKTILAHG-ITKTLHYLTLMKESALATIMNLGIEE 119

Db 281 DYIEPIKNDAAIRNYGIELAVRLCRELDSLVGLHFYILNREVAIMEVLKQLGMMTE 340

QY 120 SKVSRSLPWRBRPAPFRYKEDVRPIFWANRPKSYISRTIGMDQYPHGMGSCSNPGAL 179

Db 341 DP-RRLPALSAHPKRRREDVRPIFWASRPKSYIYRTQDMDEFPNGMGNSSPAFCGL 399

QY 180 SDYQ--FMPPRRADKLVEEMAVPLKSVEDIYERRLYCLG-----KLSNPSELDG 230

Db 400 KDYYLFYLYKSPREELTKMGEEELTSESVYEVEHNLGSPNRRGYRVCLPLMND-EP 458

QY 231 LQETRIINEOLEKINTKGLFINSOPAVNGEKSPPYGMGPGGYOKAYVEFFCSK 290

Db 459 LAAETSLKMEELLRRNRIGILITINSQPINAKPSSDPVYGMGPGSGYFQKAYLEFFTSR 518

QY 291 EKLDALVDKCD-RTSLYMAVANKDSKMSNVQTDVNAVTVGWVPAKEIIQPTIVDPS 349

Db 519 ETVEALLQVLTQYELRVNHYIDVAGENTNAPELQPNNAVTVGIPRGREIIQPTIVDPS 578

QY 350 FFWKDEAFELISRGWASLYPDEASRKLYVEGSHLYSLVNDY-INGDLFAYFAD 407

Db 579 FFWKDEAFALMIERMGKLYEESPSRTIIQYIHNDNYFLVNLVNDPELNDCLMQVED 637

RESULT 4

MTNR_SCHPO STANDARD; PRT; 603 AA.

ID MTNR_SCHPO

AC Q10258;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Probable methylenetetrahydrofolate reductase 1 (EC 1.5.1.20).

GN SPAC56P8.10

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A., Sgourou K., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tvey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Roben J., Grymoprez B., Weltjens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hubert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Puntelli B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaude V., Mottler S., Galibert F., Aves S.-J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Galliardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Rappaport G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-methylenetetrahydrofolate + NADPH.

CC -1- COFACTOR: FAD (BY SIMILARITY).

CC -1- PATHWAY: Folate metabolism.


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97051592; PubMed=8896269;
RA Ticon B., Rodriguez-Torres A.M., Rodriguez-Belmonte E., Cadahia J.L.,
RA Cerdan E.,
RT Identification of a putative methylenetetrahydrofolate reductase by
RT sequence analysis of a 6.8 kb DNA fragment of yeast chromosome VII.
RL Yeast 12:1047-1051(1996).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC methylenetetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (By SIMILARITY).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z72647; CAA96833.1; -.
DR EMU: X94106; CAA63833.1; -.
DR HSSP: P00394; 1B5T.
DR SGP: S0003093; MET13.
DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003171; Methydrf_redctase.
DR Pfam: PF02219; MTHFR_1.
DR TIGRFAMS: TIGR00677; fadh2_euk.1.
DR Oxidoreductase: Flavoprotein; FAD; NADP.
KT CONFLICT 73 73 A -> R (IN REF. 1; CAA63833).
FT SEQUENCE 599 AA: 68472 MW: AD3465B52A4E700 CMC64;
SQ

```

```

Query Match      35.6%; Score 778; DB 1; Length 599;
Best Local Similarity 40.9%; Pred. No. 1,1e-55;
Matches 168; Conservative 76; Mismatches 147; Indels 20; Gaps 11;

```

```

QY 4 IYQQLFYDIDILKLVNDCROGICPIYVGIIMPINNNKGFIRMGFCFKTRPADIMAL 63
DB 185 IYQQLFYDIDILKLVNDCROGICPIYVGIIMPINNNKGFIRMGFCFKTRPADIMAL 244
QY 64 EPIKDNEAVKAYGILGEMCKKILAHG-ITLHLTYLTMKESALAILMNLGLI-ESK 121
DB 245 DPKDDDELVRDGLNLIYEMOKLLDSGVSHHITTMNLEKAPLMIERLNIIPTESE 304
QY 122 VSRS---LPWRRPNAVFEKEDVREPIEMANRPKYSITRTIGW--DQYPHGWDSCNPS 175
DB 305 FNAHPLAVLPWRKSLNPKRKNNEVRIPIFWKRRRYSYVARTSQWAVDEPFGNRFDSSSPA 364
QY 176 YGLA-----SDYOFMRRADDKLVEMNAVPLKSVEDIYRFRFLYCLGKRSNPSGLDGL 231
DB 365 FGLDLDGSD--LIRSA--NRCLELWSP--TSINDVAFVLYNLYLGNLICKLPWSDIP-I 418
QY 232 QPRTKILNLEKINFKGLTINSOPAVNGEKSDSEPTVGWGGPGGVYVOKAVVEFGCSKE 291
DB 419 NDEINPKAKHLIELNOSHITITINSOPVAGISNCKIHMGKDKDGVVYVOKLEFMLPRT 478
QY 292 KLDALVKKCRDRTSLTYMAVNNKDGSKSN-VGQTDVNAVTVGVPAKEIIQPTIYDVPSF 350
DB 479 KLPKRLIDTLKNNFLLFYAIDSGDLLSHNPNKSNKSNVATWGIFPERELIQTPIYKISF 538
QY 351 NWKDAFAELWISKGNASLIPDEDA-SRKIVEEVGSHPFLYSIVNDNYINGD 400
DB 539 LAKKEFYHLINEMKLMNKKYDKPHSAQPLISLDYICLVNIYVNDYISP 589

```

```

RESULT 7
MTHS_SCHPO STANDARD: PRT: 641 AA.
AC 074927; P78770; Q9UT80;
DT 15-JUL-1999 (Rel. 38, Created)

```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylenetetrahydrofolate reductase 2 (EC 1.5.1.20).
GN MTHFR2 OR SPAC343.10.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968;
RA Naula N.;
RL Thesis (1998), University of Bern, Switzerland.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rayndream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Vanckereit G., Aert R., Robben J., Grymompiez B.,
RA Welljens I., Volckreels E., Rieger M., Schaefer M., Mueller Auer S.,
RA Gabell C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehner H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
RA Eger P., Cadieu E., Dreano S., Gloux S., Lelaure V., Motter S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE OF 11-641 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=9816722; PubMed=9501991;
RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs."
RL DNA Res. 4:363-369(1997).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC methylenetetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (By SIMILARITY).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ011686; CAA09738.1; -.
DR EMBL: AL109739; CAB52273.1; -.
DR EMBL: D89118; BAA13780.1; -.
DR HSSP: P00394; 1B5T.
DR InterPro: IPR004621; Fadh2_euk.
DR InterPro: IPR003171; Methydrf_redctase.

```



```

DR Pfam; PF02219; MTHFR; 1.
DR TIGRFAMs; TIGR00677; fadh2_euk; 1.
KW Oxidoreductase; Flavoprotein; FAD; NADP.
FT CONFLICT 202 202 E -> D (IN REF. 3).
FT CONFLICT 281 281 G -> A (IN REF. 3).
FT CONFLICT 455 455 A -> V (IN REF. 3).
FT CONFLICT 537 641 QVTVYAGNNSEELTNPAPKQASAVTVGYPGRREITSTII
AEVSEKAMLSSEFQWGMEMANLYSKNPSKRLLENCINDRM
LVTVIHDDPMKNGKLVLELDF -> HVRNFSSTSDST
CFORVSKPLPIIRARDRKFSVQK (IN REF. 3).
FT CONFLICT
FT SEQUENCE 614 614 I -> L (IN REF. 1).
SO SEQUENCE 641 AA; 72140 MW; E7AFDDJF2CIC8H4 CRC64;

Query Match
Best Local Similarity 32.2%; Score 629; DB 1; Length 641;
Matches 147; Conservative 80; Mismatches 161; Indels 68; Gaps 13;

OY 1 ARVIVTQLFVDTDFLFKFNVDRCRQIGTC---PIVGPIMINNYKGFIRMTGCKTKRIP 56
D 184 ADFITITQIFVEPEAFIKFENFVRNHSNALRNPIPAIMEIQSYGLKRMTRLCGCSVP 243
OY 57 ADIMALEPDKNEAVKAYGIIHCTEMCKIL-----AIGITLHLYTNMKSALAI 110
D 244 SLMQRLNAAPDDEALKNIGVEHLVDMIKKIMNVQGRVHG---FHECTLNERVALI 300
OY 111 LMLGLI-----EESKVSRS-----LPMRRP-----ANVFR 136
D 301 LKNSGLLTKRKQVSEMEDERKLTTRKSLDEPAELHNOVYVPSQAPADSSNLIFV 360
OY 137 VKE-----DVRPIFWANRPKSYISRTIGMDQYPRGRWGDCSNPSYGAISDY-- 182
D 361 TSKQSSVSGHKNDLLEAP-FVSESGSGVLGRQANWMDFTNGRRGRDPSPAYGEIDGYP 419
OY 183 QEMPRARADKTLVEKMANVPLKSVEDYIRFRFLYCLGKLRSPNSELDELQIETKTIINDQL 242
D 420 TLHPFPSSALKL--WGYPVE-SDITSLEFKHIMSDISALPMD-EPVEVEYETIKAYL 474
OY 243 EKINTKGFLLTNSQPAVANGKSDSPYVGMGGPGGYVYOKAVEFFCSKEKIDALVDKCD 302
D 475 LKLNQNSMWTYGSQPAVANGASADYVFGMGKGRVPGKAVPECFVANKDKLDITTKMD 534
OY 303 RLSLTLYAVNKDGSWKSXNVGQTDVNAVTVGPAKEIIIOPTIVDPSEFNWKADEFIWS 362
D 535 NPOVTVYVAGNNKSEFLTNAPKDGASAVTVGPREIIITSTIIAEVSEPKAWLSSEFQWG 594
OY 363 RGNASTLYPEDEASRKLVEEGSGHFLVSLVNDYIN 398
D 595 E-WANLYSKNTPSKRLLENCINDRMVLTVIHDDFMD 629

RESULT 8
MTHR_YEAST
ID MTHR_YEAST STANDARD; PRT: 657 AA.
AC P46151;
DT 01-NOV-1995 (Ref. 32, Created)
DT 01-OCT-1996 (Ref. 34, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE Methylenetetrahydrofolate reductase 1 (Ec 1.5.1.20).
GN MET12 OR YPL023C OR LIP8C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Arujo R., Aparicio A., Bartrell B.G., Badcock K., Beres V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duestelhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,

```

```

RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunkle-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,
RA Komp C., Kundl O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirzapati S., Moestl D.,
RA Mueller-Auer S., Namath A., Neutwich U., Oetner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schater M., Scharle M.,
RA Scherrens B., Schramm S., Schroeder M., Slicu A.M., Tetteijn H.,
RA Ureterazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Welsh S.V., Wambut R., Wang Y., Wedler E., Wedler H., Minnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.,
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI,"
RL Nature 387:103-105(1997).
RN [2]
RP SEQUENCE OF 1-131 FROM N.A.
RX MEDLINE=85061207; PubMed=6095044;
RA Yang E., Friedberg E.C.;
RT "Molecular cloning and nucleotide sequence analysis of the
RT Saccharomyces cerevisiae Rad1 gene."
RL Mol. Cell. Biol. 4:2161-2169(1984).
RN [3]
RP IDENTIFICATION
RX MEDLINE=95004587; PubMed=7920641;
RA Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S.,
RA Mathews R.G., Rozen R.;
RT Human methylenetetrahydrofolate reductase: isolation of cDNA,
RT mapping and mutation identification."
RL Nat. Genet. 7:195-200(1994).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADP(+) = 5,10-
CC methylenetetrahydrofolate + NADPH.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Folate metabolism.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: U36624; AAB68164.1; -
CC DR EMBL: K02070; -; NOT_ANNOTATED_CDS.
CC DR HSSP: P00394; 1B5T.
CC DR SCD, S0005944; MET12.
CC DR InterPro: IPR004621; fadh2_euk.
CC DR InterPro: IPR003171; Methydrif_redctse.
CC DR Pfam: PF02219; MTHFR; 1.
CC KW Oxidoreductase; Flavoprotein; FAD; NADP.
FT CONFLICT 110 111 DP -> NL (IN REF. 2).
FT FT 116 ED -> VV (IN REF. 2).
FT CONFLICT 119 119 ED -> V (IN REF. 2).
FT FT 131 ESPFKYAV -> RLINMRLE (IN REF. 2).
SO SEQUENCE 657 AA; 73941 MW; D75612EBA92D67500 CRC64;

Query Match
Best Local Similarity 26.5%; Score 578.5; DB 1; Length 657;
Matches 151; Conservative 80; Mismatches 155; Indels 93; Gaps 16;

OY 1 ARVIVTQLFVDTDFLFKFNVDRCRQ-IGTCPIYVGPIMPINNYKGFIRMTGCKTKRIPADI 59
D 183 ADFITITQIFVEPEAFIKFENFVRNHSNALRNPIPAIMEIQSYGLKRMTRLCGCSVP 242
OY 60 MAALP-IKDNEAVKAYGIIHCTEMCKI--LAHGITLHLYTNMKSALATIM--- 112
D 243 LSRFPPIQSDDNAVKSIGVDILIEIQTSGSRKIGFHHYTLNLEKALAOIVSQSP 302
OY 113 -----NLGLI-----EESKVSRLPMRRP 134
D 303 VLSHIYNSSEGEDETSIGSIEENVPIEDADGDIVLSDNSMETVANRRK--RRHSSL 360

```



```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF OR H11444.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: 5-methylenetetrahydrofolate + acceptor = 5,10-
CC -1- methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U32823; AAC23094.1; -.
DR HSSP: P00394; 1B5T.
DR TIGR: H11444; -.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylot_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;
KW Complete proteome.
SQ SEQUENCE 292 AA; 33020 MW; C6DCC9BC80CF994 CRC64;
-----
Query Match 8.6%; Score 189; DB 1; Length 292;
Best Local Similarity 32.8%; Pred. No. 3.6e-08;
Matches 38; Conservative 23; Mismatches 55; Indels 0; Gaps 0;
-----
OY 1 ARVITOLFVDFDIFLKVNDQIGITCPIVPGIMPINNYKGIKMTGFGCKTKIPADIM 60
DB 175 ANHVITOFFPIENTLRDRCSAGIDTEVTPILPTNKROQKMASTFNVPKIPAMLV 234
OY 61 AALEPIKNEBAVKAYGHIHGTCEMKKIIAHGIKTLHLTYLNMEKSAIATIMNGL 116
DB 235 KAYDGLDNDPTTRNLVAASVAMDMVKILSRGVNDFHYITLNRSELYALICHTLGV 290
-----
RESULT 12
METF_ECOLI STANDARD; PRT; 296 AA.
AC P00394;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF OR B3941.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=84041480; PubMed=6356036;
RA Saint-Girons I., Duchange N., Zakin M.M., Park I., Margarita D.,
RA Ferrara P., Cohen G.N.;
RT "Nucleotide sequence of metF, the E. coli structural gene for 5-10
RT methylene tetrahydrofolate reductase and of its control region.";
RL Nucleic Acids Res. 11:6723-6732(1983).
-----
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=93347969; PubMed=8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes."
RL Nucleic Acids Res. 21:3391-3398(1993).
-----
[3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=99215588; PubMed=10201405;
RA Guenther B.D., Sheppard C.A., Tran P., Rozen R., Matthews R.G.,
RA Ludwig M.L.;
RT "The structure and properties of methylenetetrahydrofolate reductase
RT from Escherichia coli suggest how folate ameliorates human
RT hyperhomocysteinemia.";
RL Nat. Struct. Biol. 6:359-365(1999).
CC -1- CATALYTIC ACTIVITY: 5-methylenetetrahydrofolate + acceptor = 5,10-
CC -1- methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: V01502; CAA24747.1; -.
DR EMBL: U19201; AAB03073.1; -.
DR EMBL: AE000468; AAC76923.1; -.
DR PIR: A00462; RDECMH.
DR PIR: S40884; S40884.
DR PDB: 1B5T; 2O-JAN-99.
DR Ecogene: EG10585; metF.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylot_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;
KW 3D-structure; Complete proteome.
SQ SEQUENCE 296 AA; 33102 MW; B702702D2BE9521E CRC64;
-----
Query Match 8.6%; Score 189; DB 1; Length 296;
Best Local Similarity 32.1%; Pred. No. 3.7e-08;
Matches 36; Conservative 25; Mismatches 51; Indels 0; Gaps 0;
-----
OY 5 VVQLFVDFDIFLKVNDQIGITCPIVPGIMPINNYKGIKMTGFGCKTKIPADIMALE 64
DB 181 ITQFFEDVESYLNKRDRCVSAGIDVEILPGILPVSNFKQAKKFAADTVNRIPAMMAOWED 240
OY 65 PINDNEBAVKAYGHIHGTCEMKKIIAHGIKTLHLTYLNMEKSAIATIMNGL 116
DB 241 GLDDDAETKLVGANIAMDMVKILSRGVNDFHYITLNRSELYALICHTLGV 292

```

```

RESULT 13
METF_SALTY STANDARD; PRT; 296 AA.
ID METF_SALTY STANDARD; PRT; 296 AA.
AC P11003;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=88302115; PubMed=2841568;
RA Stauffer G.V., Stauffer L.T.;
RT "Cloning and nucleotide sequence of the Salmonella typhimurium LT2
RT metf gene and its homology with the corresponding sequence of
RT Escherichia coli.";
RL Mol. Gen. Genet. 212:246-251(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X07689; CAA30531.1; -.
DR EMBL: AE008891; AML22945.1; -.
DR PTR: S03169; S03169.
DR HSSP: P00394; 1B5T.
DR StyGene: SG10224; metF.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylcrof_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis;
KW Complete proteome.
FT CONFLICT 94 94 A -> P (IN REF. 1).
SQ SEQUENCE 296 AA; 33173 MW; 2B834F880B56A643 CRC64;

Query Match 8.4%; Score 184; DB 1; Length 296;
Best Local Similarity 31.2%; Pred. No. 9.4e-08;
Matches 35; Conservative 26; Mismatches 51; Indels 0; Gaps 0;

QY 5 VTQLEFYDIDFLKFNVDRCQIGITCPVGPIMPINNYKGFIRMTGFCRTKIPADIMAALE 64
DB 181 ITQFFDVESYLRFRDRCVAGSAGDVEIPIGILVSNFKQAKKRAADMTNVRIPSSWMSIMRE 240
OY 65 PIKDNEAVKAYGIHLGTCKMKIILAHGKIKTLHLTYLNMKSALAILMNLGL 116

```

```

DB 241 GLDNDAEIRKLIVGANIAMDWKILSRGVKDFHRYTLNRAEMSYAICHTLGV 292
: : : | | | : : | : | | | | : : | |
RESULT 14
METF_ERWCA STANDARD; PRT; 298 AA.
ID METF_ERWCA STANDARD; PRT; 298 AA.
AC P71319;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=71;
RA Calcutt M.J., Lewis M.S., Eisenstark A.;
RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + acceptor = 5,10-
CC methylenetetrahydrofolate + reduced acceptor.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: Methionine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE
CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U74302; AAC72242.1; -.
DR HSSP: P00394; 1B5T.
DR InterPro: IPR004620; Fadh2_bact.
DR InterPro: IPR003171; Methylcrof_redctse.
DR Pfam: PF02219; MTHFR; 1.
DR TIGRFAMs: TIGR00676; fadh2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Methionine biosynthesis.
SQ SEQUENCE 298 AA; 33589 MW; 50FE729E4E3E2C56 CRC64;

Query Match 8.0%; Score 176; DB 1; Length 298;
Best Local Similarity 30.4%; Pred. No. 4.3e-07;
Matches 34; Conservative 25; Mismatches 53; Indels 0; Gaps 0;

QY 5 VTQLEFYDIDFLKFNVDRCQIGITCPVGPIMPINNYKGFIRMTGFCRTKIPADIMAALE 64
DB 181 ITQFFDVESYLRFRDRCVAGSAGDVEIPIGILVSNFKQAKKRAADMTNVRIPSSWMSIMRE 240
OY 65 PIKDNEAVKAYGIHLGTCKMKIILAHGKIKTLHLTYLNMKSALAILMNLGL 116
DB 241 GLDNDPEIRKRVKVASIAMDVKIILSRGVKDFHRYTLNRAELSYAICHTLGV 292
: : : | | | : : | : | | | | : : | |
RESULT 15
METF_BUCAI STANDARD; PRT; 292 AA.
ID METF_BUCAI STANDARD; PRT; 292 AA.
AC P57154;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5,10-methylenetetrahydrofolate reductase (EC 1.7.99.5).
GN METF OR B0046.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]

```


SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 2/24/03
Searcher: P. Schreiber
Terminal time: 18
Elapsed time: 11
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site

____ STIC
☒ CM-1 6A03
____ Pre-S

Type of Search

2 N.A. Sequence
1 A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
☒ Other Compass

